JOURNAL OF VERBAL LEARNING AND VERBAL BEHAVIOR **16**, 137–150 (1977)

Letter and Configuration Information in Word Identification

JAMES L. MCCLELLAND

University of California, San Diego

Subjects learned meanings for 16 invented words (e.g., BARDREL) and practiced categorizing each word on the basis of its meaning. During learning and practice, each word appeared consistently in either script or uppercase type. Subsequently subjects categorized both versions of each word (BARDREL and *bardrel*) on the basis of meaning. On the first exposure to each word during the test phase, categorization times were about 50 msec slower for unfamiliar versions. The difference diminished rapidly, disappearing after two to six presentations of the previously unfamiliar versions, depending on prior practice. Training transferred better from uppercase to script than from script to uppercase. The results suggest that subjects relied on configuration information, especially for script words, as well as letter identity information.

Does reading a word involve recognizing a familiar visual configuration, or does it involve recognizing a particular arrangement of letters? Many authors, from Cattell (1866) to Smith (1971) and Johnson (1975), have argued that words are read by recognizing them as familiar visual configurations. On the other hand, others have adopted the view that words are read by recognizing a pattern of letters passed forward by a preliminary letter analysis process. Huey (1908) cites several early proponents of the preliminary letter analysis view. More recently, Geyer (1970), Gough (1972), Estes (1975), Henderson (1976), and McClelland (1976) have all favored preliminary letter analysis.

This research was partially supported by Grant No. PHS MH 15828 to the Center for Human Information Processing, University of California, San Diego, and by the Academic Senate of the University of California, San Diego. I would like to thank Eric Devos for running all the subjects and analyzing much of the data; I would also like to thank Lynn Cooper and Dave LaBerge for their many helpful comments. Experiment 3 grew out of a conversation with Dave. T. K. Landauer also provided many useful comments.

Requests for reprints should be sent to James L. McClelland, Department of Psychology C-009, University of California, San Diego, La Jolla, CA 92093.

One component of the evidence often cited against a pure preliminary letter analysis model of word recognition is the evidence on the role of word shape information in the identification of words. Several investigators have shown that visual information about a word's global shape or outline plays a role in word recognition. There are several relevant findings. (1) Words with distinctive outlines can be correctly identified at a viewing distance too great for recognition of single letters (Huey, 1908). (2) Briefly presented lowercase words with unique outlines can be reported correctly more often than words with common outlines, and words with common outlines (e.g., *lint*) are often incorrectly identified as other words with the same outline (e.g., *list*; Havens & Foote, 1965). (3) Information about the global shape of words can be picked up in peripheral vision, and this information can facilitate foveal processing on the next fixation (Dodge, 1907; McConkie & Rayner, 1975; Rayner, 1975).

All of these findings seem to show that we retain in memory representations of the outlines of words, and that we match the stored shape information against information extracted from the stimulus to provide at least

ISSN 0022-5371

one indication of the identity of the word. This view was apparently held by Erdmann and Dodge (see Huey, 1908, and Woodworth, 1938, for discussions). However, closer examination in light of current preliminary letter analysis-based models of word recognition reveals that a role for *stored* word shape information has not been established. The entire set of results described above actually follows from preliminary letter analysis models involving a parallel letter analysis process which need not be complete or exhaustive before results begin to pass forward to a word identification stage. Such models have recently been proposed by Estes (1975), Henderson (1975), and McClelland (1976) to account for tachistoscopic word superiority effects (Reicher, 1969) and for equal speed in responding to word and letter displays in visual matching tasks (Johnson, 1975).

To explain these "word shape" results, preliminary letter analysis models could begin with the fact that global shape information extracted from a distant, brief, or parafoveal stimulus word can be used to determine which letters could have been present in the word at each letter position. For example, the outline of the word *shape* provides enough information to determine that the first, third, and fifth letters are all in the set a, c, e, m, n, o, r, s, u, v, w, x, and z; that the second is either b, d, f, h, k, l, or t; and that the fourth is either g, j, p, q, or y. Any more detailed information extracted from a word with this shape would of course help limit the set of possible letters even more. In any event, the sets of possible letters for each position could be determined by a preliminary letter identification stage, using both global word shape information and any other more detailed information available. The results of this process could then be passed forward to a word identification process which would attempt to find a word consistent with one of the possible letters in each position. From this sort of reasoning it should be clear that word shape information need not be stored explicitly in memory for global word shape information to be an effective cue to word identity.

Recently, Coltheart and Freeman (1974) have presented a different kind of evidence for a role of word shape information in word recognition. These investigators found that tachistoscopic report of words was disrupted when the words were presented in mixed upper- and lowercase type, compared to either all uppercase or all lowercase. Other investigators (Brooks, Note 1; Fisher, 1975) have obtained similar results using other tasks (but see Smith, 1969; Smith, Lott, & Cronnel, 1969). This disruptive effect of mixing cases appears to support the view that words are recognized in part in terms of familiar visual configurations, but it is not necessary to attribute the effect of mixing cases to a disruption of word recognition per se. Instead, mixing cases appears to disrupt some stage or aspect of processing involved in the processing of both words and (at least pronounceable) nonwords. McClelland (1976) and Taylor, Miller, and Juola (Note 2, Experiment 2) both report equivalent disruptive effects of mixing cases in processing words and pronounceable nonwords in two quite different tasks.[1] Adams (Note 3) found equivalent disruptive effects in words, pronounceable nonwords, and unrelated letter strings. All of these studies found the traditional word advantage over nonwords in accuracy and speed of processing, using both same and mixed case stimuli. In sum, these mixed case findings do not provide any firm evidence that readers use stored information about the visual forms of words in reading.

Experiment 1

The experiments reported here look for a new sort of evidence on the role of stored word

[1] Taylor, Miller, and Juola (Note 2) report a larger disruptive effect of mixing cases with words than with their pseudowords in their Experiment 3, but many of the pseudowords used in that experiment were not pronounceable (e.g., TAGRN).

shape information in word recognition. Subjects were taught to recognize and define 16 invented words (e.g., BARDREL) and were then given practice accessing the meanings of the words and responding discriminatively on the basis of the meanings. After practice, subjects were tested for speed of accessing and responding to the meanings of the words in both familiar and unfamiliar versions. This test permits comparison of the speed and accuracy of processing words when their visual configurations are familiar to the reader, and when the words are familiar, but their visual configurations are not.

Method

Stimuli. Sixteen words were invented for use in the experiment. Each of the invented words was pronounceable; eight of the words were five letters long and eight were seven letters long. To ensure that all of the words would be quite different in form between their script and uppercase versions, all the words were made up mostly of letters which differ between script and uppercase. The script form of each word had at least one ascender or descender in it, to ensure that the outline differed between the two versions as well. Uppercase versions of the words were typed using a Courier 12 IBM selectric element. Script versions were typed using the lowercase letters on the 12-point script element. The actual invented words, each one typed in both of the fonts used in the experiment, are presented in Figure 1.

Two invented meanings were constructed to go with each word. One meaning always designated an animate object (e.g., a young whale) and one always designated an inanimate object (e.g., a loud bell). Each individual subject learned animate meanings for eight words and inanimate meanings for the other eight, with the assignment of meanings to words counterbalanced jointly over subjects and word lengths.

Each subject saw each word in one and only one form during the learning and practice

BARDREL	*bardrel*	DROAF	*droaf*
SHAFFAN	*shaffan*	SHANG	*shang*
BLEAFER	*bleafer*	SPEAT	*speat*
DESHERY	*deshery*	DIGHT	*dight*
GESTARD	*gestard*	GREND	*grend*
DRUBBIT	*drubbit*	DEASH	*deash*
DRAPPER	*drapper*	BRALD	*brald*
QUIDDET	*quiddet*	PRAST	*prast*

FIG. 1. The set of invented words in the two different typefaces used in the experiments.

phases of the experiment. For each subject, four animate and four inanimate words were shown in uppercase type; the other eight words were shown in script. The assignment of typefaces to words was completely crossed with subjects, meanings, and word lengths.

Subjects. Eight University of California, San Diego, undergraduates participated in the experiment for either $2 per hour or course credit.

Procedure. At the beginning of the experiment, each subject was told to memorize the assigned meaning for each of the 16 invented words. On first presentation, the experimenter showed the subject each word in the assigned typeface with its meaning typed in the same typeface and asked the subject to attempt to memorize the meaning using any elaborative techniques he wished. On subsequent presentations, the experimenter showed only the word, and the subject had to attempt to recall the head noun of the definition of the word. For example, for the meaning "a loud bell", the subject had only to recall "bell". (The adjectives were only included to increase the concreteness, and hence the memorability, of the meanings.) If the subject could not recall, or if he recalled incorrectly, the experimenter showed him the meaning again; otherwise the experimenter simply went on to the next word without comment.

This procedure continued until the subject was able to recall all the meanings correctly on two successive runs through the entire set of 16 words. An average of 14.4 runs was required to reach criterion. Between runs through the set of cards, the cards were shuffled so that there was a new random order for each run.

The subject then received six blocks of informal categorization practice. In this practice, the experimenter simply showed each word to the subject, one word at a time, and watched the subject to see whether he tapped the table with his right index finger (to indicate that the word referred to an animate concept) or his left index finger (to indicate that it referred to an inanimate concept). The experimenter pointed out any tapping errors the subject made. Each of the six blocks of informal categorization practice included the full set of 16 words. Stimulus cards were shuffled between blocks.

After the informal practice, each subject received six blocks of formal categorization practice. The words were displayed in a two-field tachistoscope at a distance of 39 cm. Five-letter words subtended 1.8° and seven-letter words subtended 2.5°. Each display was preceded and followed by a white field with four dots framing the location of the display word. Each display lasted 200 msec, quite long enough for errorless identification. Instead of tapping, the subject pressed a response key with his right or left index finger to make his categorization responses. After six such blocks the first session ended.

Session 2 began about 24 hours later with six blocks of informal practice, followed by six blocks of formal practice. After the end of these 12 practice blocks, the experimenter informed the subject that the words could henceforth be presented in either the familiar version the subject had already seen or in the unfamiliar version which had not been exposed previously. Immediately after giving the subject this information, the experimenter gave the ready signal for the first trial of the test phase.

In the test phase of the experiment, each subject viewed both the familiar and the unfamiliar version of each word a total of three times, once in each of three successive blocks. The 32 stimuli were divided into two subgroups of 16, with each subgroup containing both versions of eight of the words. The subgroups were balanced for learned form and for number of words with animate and inanimate meanings, and subgroups were presented in alternation with order counterbalanced over subjects. The order of cards in each subgroup was randomized before each block with the restriction that the two versions of the same word could not occur within four trials of each other. Each test block of the experiment represents a fully counterbalanced design, with each word appearing exactly twice over the eight subjects in each combination of learned meaning (animate vs. inanimate) tested version (script vs. uppercase) and tested version familiarity (familiar vs. unfamiliar).

Results

Mean reaction times to categorize the familiar and unfamiliar versions of the 16 words are shown for each of the three stimulus exposures during the test phase in Figure 2. In addition, baseline reaction times are shown for the mean of the last three practice blocks

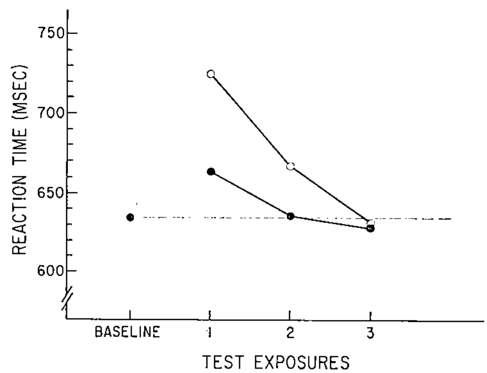

FIG. 2. Reaction time to classify familiar versions (filled dots) and unfamiliar versions (unfilled dots) of the 16 words for the last three practice exposures (baseline) and for the first, second, and third exposures of words during the test phase. Data are from Experiment 1 in which the test phase of the experiment followed the practice phase with no delay.

TABLE 1

PERCENT ERRORS

	Baseline	Test trials 1	2	3	4–6	7–18
Experiment 1						
Familiar	2.4	3.1	1.6	0		
Unfamiliar		3.9	2.3	2.3		
Experiment 2 (delayed test)						
Familiar	1.6	3.1	1.6	0.8		
Unfamiliar		6.3	0	2.3		
Experiment 3 (extended practice)						
Familiar	1.2	1.6	2.9	2.3	3.8	2.7
Unfamiliar		6.6	4.7	5.1	4.4	4.8

before the beginning of the test phase. Error trials and trials with extra long reaction times (1.5 seconds or greater) are not included in the reaction time (RT) data presented in Figure 2; percentages of errors are indicated in Table 1. Long reaction times occurred on 1.8% of trials in the test phase. More long RT's occurred for unfamiliar versions, and the rate decreased over blocks for both familiar and unfamiliar versions.

Inspection of Figure 2 reveals that mean reaction time on the first presentation of the unfamiliar versions of the words was 61 msec slower than mean reaction time to the familiar versions during the same block of trials. It further appears that reaction times to the familiar versions of the words were slowed 30 msec compared to baseline values. After the first test block, the familiar–unfamiliar difference diminished rapidly, disappearing completely by the third test block. Performance on the familiar versions of the words was back down to baseline levels by that time as well.

Statistical analyses performed over subjects reveal that the familiar–unfamiliar difference on the first trial was reliable, $t_D\ (7) = 2.36$, $p < .05$, $MS_e = 5{,}345$. Analyses are not reported over stimuli since the generality of the experiment to the population of invented words is limited by the nonrandom nature of the stimulus generation procedure. However, it is worth noting that the familiar form advantage was in the right direction on the first trial for all 16 of the words. In the second block of trials, the reaction time difference between familiar and unfamiliar versions was marginally reliable, $t_D\ (7) = 1.72$, $.1 < p < .05$, $MS_e = 2{,}958$, and on the third block there was no indication of a statistically reliable difference, $t_D\ (7) = .22$, $MS_e = 2{,}521$. Errors (Table 1) and long RTs were too infrequent for meaningful statistical analyses.

Discussion

The subjects in Experiment 1 clearly learned something specific to the version of each word that they saw during learning and practice. An experiment by Hintzman and Summers (1973) shows a similar effect in a recognition memory paradigm. At the same time, Ex periment 1 demonstrates considerable transfer of learning to the unfamiliar versions of the words. Performance on unfamiliar versions in the first test block was somewhat slower but nearly as accurate as performance on familiar versions, and only two exposures to the unfamiliar versions of the stimuli were required to bring reaction time to the level reached during practice on the familiar versions of the stimuli.

EXPERIMENT 2

One interpretation of the results of Experiment 1 could be that the disappearance of the familiar version advantage was simply due to the passage of time. Some relatively transient form-specific traces which decayed over the few minutes required for the first test presentation could account for all the data. This hypothesis predicts that the familiar version advantage would not be found in a delayed test. Experiment 2 provides a test of this prediction.

Method

Experiment 2 was identical to Experiment 1 in all respects, except that the subjects had a 1-day delay between practice and test. In the test session, each subject was immediately introduced to the test phase without any warm-up. A new group of eight subjects was used. These subjects required an average of 16.2 runs through the words during initial learning.

Results

Reaction time to categorize both the familiar and the unfamiliar versions of each stimulus appeared to be slowed a bit by the 1-day delay between practice and test (Figure 3). Aside from this effect, the results of Experiment 2 were not much different from the results of Experiment 1. The familiar versions were processed 40 msec quicker than the unfamiliar versions on the first trial, $t_D(7) = 2.44$, $p < .05$, $MS_e = 2{,}812$. As in Experiment 1, the difference was diminished in size on the second trial, although it was still reliable, $t_D(7) = 2.47$, $p < .05$, $MS_e = 1{,}076$. Again, as in Experiment 1, the difference was gone by the third trial, $t_D(7) = .05$, $MS_e = 1{,}178$, and reaction times to both familiar and unfamiliar versions were down below baseline levels. Errors and long RTs (Table 1) seem to follow the same trends as the RT data but again were too few for meaningful analysis.

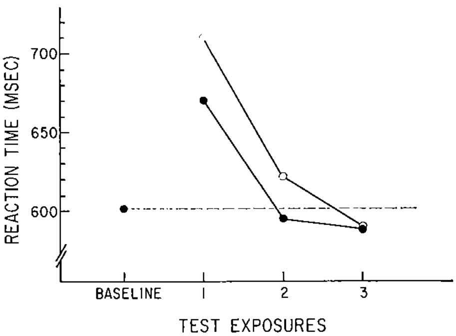

FIG. 3. Reaction times for familiar (filled dots) and unfamiliar (open dots) versions of the stimuli for each test exposure in Experiment 2, in which there was a 1-day delay between practice (baseline) and test.

Comparison of Results of Experiments 1 and 2

An analysis of variance was performed to determine whether there was any effect of the delay on the size of the *difference* in reaction time between familiar and unfamiliar versions of the words. The analysis affirmed the decreasing trend over blocks, $F(2, 14) = 3.80$, $p < .05$, $MS_e = 5{,}022$ but neither the effect of delay condition, $F < 1$, $MS_e = 4{,}180$, nor the interaction of block by delay, $F < 1$, $MS_e = 5{,}022$ were reliable.

Discussion

Although a sympathetic eye might discern a hint of a reduction of the size of the form familiarity effect in Experiment 2, statistical analyses revealed no reliable difference between the two experiments in terms of the magnitude of the effect of form familiarity. Whether or not a real diminution in the size of the effect escaped detection, it is clear that the bulk of the familiar–unfamiliar difference obtained in Experiment 1 was replicated with a 1-day delay between the last practice trial and the first test trial. This replication demonstrates that the pattern of results obtained in Experiment 1 was not due simply to transient after-effects of processing.

EXPERIMENT 3

In both Experiments 1 and 2, only two

exposures to the previously unfamiliar versions of the words were needed to complete acquisition of processing efficiency equal to the efficiency acquired during about 40 exposures to the familiar versions of the words during learning and practice.

The rapid acquisition of 40 exposures worth of processing efficiency for the previously unfamiliar versions might be taken as a reflection of the ease of acquiring form-specific processing skills. On the other hand, this rapid acquisition might be taken as a reflection of the relative unfamiliarity of the supposedly familiar versions of the words. The average of 40 exposures during learning and practice in Experiments 1 and 2 is clearly much smaller than the number of exposures mature readers generally experience for familiar words such as *the*, *and*, *it*, *more*, *this*, and many others. While it is beyond the reach of the present investigation to extend practice into a range comparable with the practice a mature reader has had processing very frequent words, it was of interest to determine whether further practice might result in the acquisition of the ability to process familiar versions of words in a manner which could not be acquired in just a very few exposures. To get evidence on this issue, Experiment 3 tripled the amount of practice on the familiar versions of the stimuli before the beginning of the test phase.

Method

The experiment once again repeated the method of Experiment 1 except that subjects received 72 blocks of practice trials spread out over four daily sessions. Session 1 was identical to the first session in Experiments 1 and 2. Session 2 began as in the previous experiments with 6 blocks of informal practice and 6 blocks of formal practice. After these 12 blocks there were an additional 6 informal and 6 formal practice blocks. The same sequence was repeated on Day 3. On Day 4, after 6 informal and 6 formal practice blocks, the test phase began just as in Experiment 1. This time, there were 6 test blocks instead of 3 as in Experiments 1 and 2. After a first group of eight subjects were run, preliminary data analysis revealed that a slight familiar version advantage was still present at the end of 6 test blocks. To determine how long the effect would last, a second group of eight subjects was run. These subjects were given 12 additional test blocks 24 hours after the first 6. The 16 subjects required an average of 17.1 runs through the deck during original learning.

Results

The additional practice produced reliably faster responding at the end of practice by the 16 subjects in Experiment 3 (Figure 4) as compared to the final practice reaction times

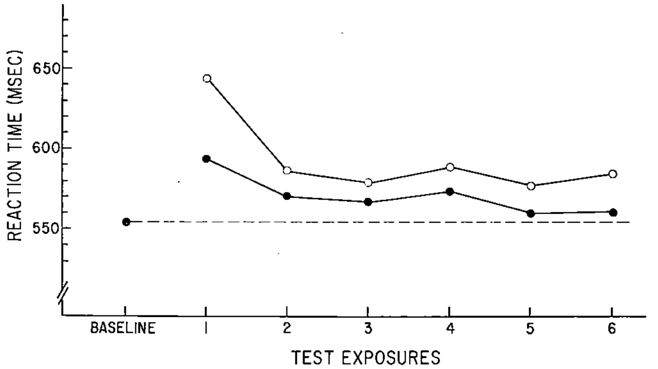

FIG. 4. Reaction time data for Experiment 3 in which the subjects received extended practice before the beginning of the test phase, for familiar (filled dots) and unfamiliar (open dots) versions of the stimuli.

from subjects in Experiments 1 and 2. Mean reaction time for the last three blocks of practice was 554 msec, compared to 617 msec average for Experiments 1 and 2, $t(30) = 2.07$, $p < .05, MS_e = 14{,}820$.

The results for the first six test blocks revealed an initial familiar version advantage quite similar to that observed in Experiments 1 and 2, although here a small effect was still present at and beyond the third test exposure to each word. For the first test exposure, the 50-msec difference between reaction times to categorize familiar and unfamiliar versions of the words was highly reliable, $t_D\ (15) = 4.39$, $p < .001$, $MS_e = 2{,}085$. Thereafter, the difference appeared relatively constant, averaging 17 msec. An analysis of the data from blocks 2 through 6 revealed a reliable form familiarity effect, $F(1, 15) = 40.55, p < .001, MS_e = 2{,}215$ but no reliable decreases in reaction time over blocks, $F(1, 60) < 1$, $MS_e = 6{,}651$, no residual blocks effect, $F(3, 60) < 1$, $MS_e = 6{,}651$, no reliable decrease in the size of the effect of familiarity over blocks, $F < 1$, $MS_e = 4{,}016$ and no residual blocks by familiarity interaction, $F < 1$, $MS_e = 4{,}016$. As in previous experiments, error rates (Table 1) and long reaction times seemed to mirror the trends present in the reaction-time data, but again were too few for meaningful analysis.

Figure 5 presents reaction-time results for blocks 7 through 18 for those eight subjects who participated in the additional blocks. Although there was a 5.5-msec difference between familiar and unfamiliar versions in the data from these blocks, this difference was neither reliable over subjects, $t_D\ (7) = 1.18$, $p > .1$, $MS_e = 2{,}167$, nor consistent over blocks. Furthermore, average performance dropped below baseline for both familiar and unfamiliar versions. A tiny lingering indication of a familiarity effect was present in the error-rate data. The 4.9% error rate for unfamiliar versions was reliably greater than the 2.7% error rate for familiar versions, t_D $(7) = 2.95, p < .05$.

Font and Familiarity

The results presented thus far have collapsed the data over the two different physical versions of the words. These analyses do not indicate whether the familiar version advantage was equally large and durable for uppercase and script stimuli, nor do they indicate whether the disruption of the processing of familiar versions of the words at the beginning of the test phase was equally large and durable in script and uppercase versions of the words.

To provide sufficient evidence to make a meaningful comparison of the size of the various effects obtained in script and uppercase stimuli for the first test exposure results, the data from all 32 subjects who participated in the three different experiments were pooled. To look for the joint effects of font and familiarity in the persisting familiar version

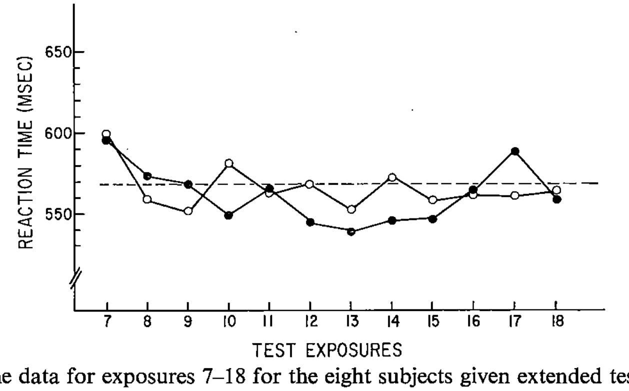

FIG. 5. Reaction time data for exposures 7–18 for the eight subjects given extended testing in Experiment 3, for familiar (filled dots) and unfamiliar (open dots) versions of the words. Note condensation of the abscissa. Dotted line is the baseline for these eight subjects.

advantage obtained in exposures 2 through 6 after extended practice, the data from the 16 subjects who participated in Experiment 3 were considered. Finally, the second test session data from the eight subjects in Experiment 3 who received extended testing were analyzed to give some idea of the final processing of script and uppercase words after the familiar version advantage had virtually disappeared.

The relevant results, collapsed over subjects (and experiments where multiple experiments are included in the same analysis), are presented in Table 2.

Effects of font of familiar versions of words during baseline and test phases. The pooled baseline and first test exposure data revealed slightly faster responding to the familiar uppercase words. Also, there was some disruption of processing familiar versions of the words in both script and uppercase at the beginning of the test phase. Analysis of variance revealed a main effect of font, $F(1, 30) = 4.84$, $p < .05$, $MS_e = 2{,}439$ and a main effect of condition (baseline vs. test), $F(1, 30) = 22.36$, $p < .01$, $MS_e = 2{,}943$. The interaction was not reliable, $F < 1$, $MS_e = 1{,}645$.

The results from exposures 2–6 for the 16 subjects in Experiment 3 and the results from exposures 7–18 for the 8 subjects who received extended testing again suggest a slight reaction-time advantage for familiar versions of uppercase words over familiar versions of script words. The 16-msec uppercase advantage was reliable in the exposure 2 to 6 data, $F(1, 15) = 6.49$, $p < .05$, $MS_e = 277$. The 10-msec uppercase advantage in the exposure 7 through 18 data did not reach reliability, $F(1, 7) = 1.82$, $p > .1$, $MS_e = 439$, although six of the eight subjects showed differences in the expected direction.

Effects of font on the familiar version advantage. Practice reading a word in uppercase

TABLE 2

REACTION TIME BY FONT OF TESTED ITEM FOR WORDS LEARNED IN THE SAME FONT OR IN THE OTHER FONT[a]

	Item tested in	
	Script	Uppercase
Baseline[b]	594 (1.3)	578 (1.8)
Test Presentation 1[b]		
Same font learned	642 (2.0)	620 (2.7)
Other font learned	676 (7.4)	684 (4.3)
Transfer decrement	36	64
Test Presentations 2–6[c]		
Same font learned	574 (3.6)	558 (3.3)
Other font learned	578 (4.7)	588 (4.1)
Transfer decrement	4	31
Test Presentations 7–18[d]		
Same font learned	565 (2.3)	555 (3.0)
Other font learned	567 (4.8)	564 (4.8)
Transfer decrement	2	9

[a] Numbers in parentheses are percentage of errors.
[b] $n = 32$.
[c] $n = 16$.
[d] $n = 8$.

transferred better to reading the word in script than practice reading a word in script transferred to reading the uppercase version. On the first presentation, reaction times for words learned in uppercase and tested in script were only 34 msec slower than times for words learned and tested in script, while reaction times for words learned in script and tested in uppercase were 64 msec slower than times for words learned and tested in uppercase. Over presentations 2–6, reaction times for words learned in uppercase and tested in script were only 4 msec slower than words learned and tested in script, while times for words learned in script and tested in uppercase remained 31 msec slower than for words learned and tested in uppercase. The pattern persisted through test presentations 7 through 18, where only 2 msec separates reaction times for words learned in uppercase and tested in script from those learned and tested in script but 9 msec separates reaction times for words learned in script and tested in uppercase from words learned and tested in uppercase.

Analyses of variance produced the following results: For the first test exposure data, only the effect of familiarity, $F(1, 30) = 25.49$, $p < .001$, $MS_e = 2{,}877$, was reliable. The main effect of tested font was not reliable, $F < 1$, $MS_e = 2{,}867$, nor was the apparent interaction, $F(1, 30) = 1.37$, $p > .1$, $MS_e = 4{,}149$. For the second through sixth exposure data, the effect of familiarity was again reliable, $F(1, 15) = 20.28$, $p < .001$, $MS_e = 230$. The effect of font was not reliable, $F < 1$, $MS_e = 203$. Here, however, the quite obvious interaction was reliable, $F(1, 15) = 6.51$, $p < .05$, $MS_e = 362$. The results from exposures 7–18 for the subgroup of eight subjects given extended testing were most notable for the small size of the differences apparent in the table. Nevertheless, a marginally reliable main effect of font emerged, $F(1, 7) = 4.31$, $.1 > p > .05$, $MS_e = 773$, reflecting the fact that responses to uppercase words were slightly faster than responses to script words, collapsing over familiar and previously unfamiliar versions at this phase of testing. Neither the effect of familiarity, $F(1, 7) = 1.37$, $p > .1$, $MS_e = 2{,}167$ nor the interaction, $F < 1$, $MS_e = 2{,}708$, approached reliability, though the pattern of results appeared to be just a reduced version of the pattern in the other cells.

General Discussion

Learning to read a word in one type font (uppercase or script) transferred extensively to processing the word in the other font, but the transfer was not quite complete. On the very first exposure to unfamiliar versions of familiar words, subjects categorized the unfamiliar versions more slowly than they categorized familiar versions of the words. The size of the effect was largely independent both of the extent of practice on the familiar version of the words and of the presence or absence of a 1-day delay of testing after practice. The size of the effect was difficult to gauge exactly: Reaction times to the first exposures of unfamiliar versions averaged 85 msec longer than baseline reaction times to familiar versions, but only 50 msec longer than reaction times to familiar versions presented at the same point in the experiment as the unfamiliar versions. Depending on the meaning attached to the increase in reaction times for familiar versions of words, the advantage for familiar over unfamiliar versions was somewhere between 50 and 85 msec on the first exposure to the unfamiliar versions.

The familiar–unfamiliar reaction-time difference disappeared as it was being measured. Given about 40 previous exposures to one version of each word, the disadvantage of the previously unfamiliar version was eliminated after two exposures. Given more extended practice (about 90 exposures), the disadvantage of the unfamiliar version remained reliable until after the sixth test exposure to the previously unfamiliar version, though the disadvantage was considerably reduced in size after the very first exposure. Since the reaction times obtained with familiar versions of each word reverted to baseline levels as the familiar–unfamiliar difference disappeared, there was little ambiguity about the number of exposures needed for the familiar version advantage to disappear.

The final important findings of the study involve differences between fonts both in terms of processing familiar versions and in terms of the transfer of processing efficiency from one particular version to the other. For one thing, there was a slight but consistent advantage for familiar uppercase versions of words compared to script versions. This effect is difficult to interpret because there are a variety of uncontrolled differences between the type fonts used. More importantly, the familiar version advantage was larger for words learned in script and tested in uppercase than for words learned in uppercase and tested in script. While this effect was not reliable in the first test exposure results, a trend was apparent, and the effect was both apparent and reliable in the data from the second through sixth test exposures in Experiment 3. The pattern of results showed that learning to process a word in uppercase transfers more completely to processing the word in script than learning to process a word in script transfers to processing the word in uppercase.

It should be borne in mind that the results of these experiments may be rather specific to the exact nature of the task at hand. Indeed, it may not be the case that subjects read the words in a fashion exactly analogous to the way they would read words in a normal reading situation. Subjects may have relied on form-specific cues more than they would in a normal reading context. In normal reading, just knowing that a word is uppercase or script is very unlikely to provide much useful restriction of possible words, but in the present experimental context, at least during practice, just determining whether the word was uppercase or script would be sufficient to rule out half of the possible alternative stimuli. While the results may be somewhat task-specific, they do point toward certain conclusions about the way in which perceivers can process words, and these conclusions relate importantly to general models of word recognition.

Theoretical Implications

Direct recognition models. The familiar version advantage might be seen as support for models which stress global shapes as the usual cue to word identity (e.g., Erdmann & Dodge, discussed by Huey, 1908 and Woodworth, 1938) or for models which suggest that the basis for word recognition is the "total word picture" (Woodworth, 1938) or the results of a position-specific feature analysis (Smith, 1971). However, these models have some difficulty with two aspects of the results. First, they provide little basis in and of themselves for accounting for the fact that subjects were able to read unfamiliar versions of words at all. There are two ways for direct recognition models to deal with the finding that subjects processed unfamiliar versions only a bit more slowly than familiar versions even on the first exposure of the unfamiliar versions. First, we might postulate a backup mechanism relying on preliminary letter analysis. Alternatively, we might suggest that there is sufficient visual similarity between fonts for subjects to actually recognize the unfamiliar versions directly. Of course, if the visual similarity were dependent on learning to extract invariant features of letters in different fonts, it would be tantamount to preliminary letter analysis. The rapid disappearance of the form familiarity effect also poses difficulties. At first glance it seems tempting to suggest that subjects were able to learn to recognize words directly from their visual configurations in two trials. This suggestion could account for the present findings, but it leads to an unfortunate prediction. If subjects could learn to recognize words efficiently in just two exposures to their visual forms, then learning to read should take only a few presentations of (each of several different versions of) all the words to be learned. Unfortunately, learning to read is not so easy, and teaching reading by the "whole word" method, which earlier proponents of direct recognition advocated (see Huey, 1908, or Smith, 1971, for a discussion), only makes matters worse (Chall, 1967). Thus, the facts about learning to read seem to contradict the idea that subjects can learn to identify words directly from the visual configuration in just two exposures.

Models relying on preliminary letter analysis. The small size and rapid disappearance of the familiar version advantage are consistent with a role for preliminary letter analysis in accessing the meanings of visually presented words. Evidence of a role for preliminary letter analysis in a different kind of task (reading aloud) comes from an experiment by Brooks (Note 1). Instead of teaching subjects to read invented words in familiar fonts, Brooks taught subjects to read familiar words in invented fonts. In one condition, the letters in the new font were in one-to-one correspondence with the appropriate letters in the standard alphabet, so subjects could decode the font as a cue to the identity of the words. In another condition, the letters in the new font were not consistently paired with letters in the alphabet, so that each word had to be learned simply as a configuration paired with a name. Subjects in the correspondence condition read the words aloud faster than subjects in the paired associate condition. The effect required 200 trials to materialize, but it persisted thereafter until the experiment ended after 400 trials. That this effect ever emerged at all is particularly striking in view of the fact that subjects were not given any explicit alphabetic association training in either condition. These results of Brooks', taken together with those presented here, paint a picture of a role of alphabetic decoding in processing visually-presented words in both meaning access and vocal reading tasks.

The fact that there was a familiar version advantage at all makes it difficult to maintain a pure preliminary letter identification based model of word recognition. Proponents of such a model might be tempted to argue that the familiar version advantage is simply due to a reliance on the detection of the font of the word, as one cue to its identity, in addition to the identity of the letters in the word. However, this argument does not explain why learning to read uppercase transferred to script better than learning to read script transferred to uppercase.

A combined model. Perhaps the most reasonable model is one which states that word recognition involves some direct recognition in conjunction with preliminary letter analysis. In one version of such a model (LaBerge, Note 4) information extracted from a stimulus might be subjected to preliminary letter analysis and passed forward to word analysis at the same time that other information concerning the outline or global shape of the word would be passed to word analysis directly. One might also postulate the visual analysis of familiar letter cluster groups within such a system (cf. Landauer, Didner, & Fowlkes, Note 5; Smith, 1971).

Let us assume that experience reading a word results in connection of a detector for the word to the outputs of detectors of the letters in the word and detectors for the word's outline shape. On later reading we assume that the word detector accrues activation from these detectors until a threshold for response activation is reached. Under these assumptions, threshold would be reached most rapidly for familiar words in familiar visual forms because detection of both letter and shape information would contribute to the activation. Unfamiliar versions of words would also activate the word detector, but it would reach threshold more slowly because only the letter information would provide activation.

The rapid disappearance of the form familiarity effect does pose some difficulties for this model, unless we make the same assumption of very rapid form specific learning which we already rejected above. However, it is not unreasonable to suppose that a subtle difference in processing between familiar and unfamiliar words which is only a partial determinant of performance might not be reflected very strongly in behavior. In addition, it is possible that subjects tended to rely more heavily on letter identity cues during the test phase, particularly for words presented in script, further obscuring effects of form familiarity.

In any case, the combined model does provide one account for the fact that practice reading a word in uppercase transferred better to script than practice in script transferred to uppercase. Initially, the model would seem to imply just the opposite. Since words in script have distinctive configurational features while words in uppercase do not, we would expect subjects to learn to read script words both in terms of the letters and in terms of the configuration, and to learn to read uppercase words primarily in terms of the letters alone. From this expectation it seems to follow that practice reading words in script should transfer nearly completely to reading uppercase, while practice reading words in uppercase should transfer somewhat less well. However, it is possible that when the subjects were learning to read words in script, they did not actually learn to use the letter identity information as effectively as they did when learning to read words in uppercase, because they were able to rely on global configuration information to a greater extent with words in script. Thus, it is possible that the incomplete transfer of acquired processing ability from words in script to words in uppercase might not reflect a *form*-familiarity effect at all; it may in fact reflect a *letter*-familiarity effect, at least in part.

This last suggestion is quite speculative, and further research is required to see if it is correct. In the meantime, however, one point is worth noting. If distinctive configural cues do interfere with learning to use letter identity information in reading, there are important implications for how reading should be taught. Reliance on global shape information might be useful in those special situations, such as the practice phase of the present experiment, in which word identity is consistently cued by a distinctive global shape. However, reliance on global shape information would not be so useful in everyday life, where words can occur in any one of several different calligraphies, and the only invariant cues to the identity of words are the identities of the letters the words contain. Under these circumstances, it may well be that beginning readers will acquire reading skills of more general utility if they are taught to read words whose global shapes are as undistinguished as possible, so that maximal reliance on letter identity information will be ensured.

References

Cattell, J. M. The time taken up by cerebral operations. *Mind*, 1886, **11**, 220–240; 377–392.

Chall, J. *Learning to read: The great debate.* New York: McGraw-Hill, 1967.

Coltheart, M., & Freeman, R. Case alternation impairs word identification. *Bulletin of the Psychonomic Society*, 1974, **3**, 102–104.

Dodge, R. An experimental study of visual fixation. *Psychological Review Monograph Supplement*, 1907, **8**, 35.

Estes, W. K. The locus of inferential and perceptual processes in letter identification. *Journal of Experimental Psychology: General*, 1975, **2**, 122–145.

Fisher, D. F. Reading and visual search. *Memory and Cognition*, 1975, **3**, 188–196.

Geyer, J. J. Models of the perceptual process in reading. In H. Singer & R. Ruddell (Eds.), *Theoretical models and processes in reading.* Newark, Delaware: International Reading Association, 1970.

Gough, P. B. One second of reading. In J. F. Kavanaugh & I. V. Mattingly (Eds.), *Language by ear and by eye.* Cambridge, Massachusetts: MIT Press, 1972.

Havens, L. L., & Foote, W. E. The effect of competition on visual duration threshold and its independence of stimulus frequency. *Journal of Experimental Psychology*, 1963, **65**, 6–11.

Henderson, L. Do words conceal their component letters? A critique of Johnson (1975) on the Visual Perception of words. *Journal of Verbal Learning and Verbal Behavior*, 1975, **14**, 648–650.

Hintzman, D. L., & Summers, J. J. Long term visual traces of visually presented words. *Bulletin of the Psychonomic Society*, 1973, **1**, 325–327.

Huey, E. B. *The psychology and pedagogy of reading.* New York: Macmillan, 1908.

JOHNSON, N. F. On the function of letters in word identification: Some data and a preliminary model. *Journal of Verbal Learning and Verbal Behavior*, 1975, **14**, 17–30.

MCCLELLAND, J. L. Preliminary letter identification in the perception of words and nonwords. *Journal of Experimental Psychology: Human Perception and Performance*, 1976, **2**, 80–91.

MCCONKIE, G., & RAYNER, K. The span of effective stimulus during an eye fixation in reading. *Perception and Psychophysics*, 1975, **17**, 578–587.

RAYNER, K. The perceptual span and peripheral cues in reading. *Cognitive Psychology*, 1975, **7**, 65–81.

REICHER, G. M. Perceptual recognition as a function of meaningfulness of stimulus material. *Journal of Experimental Psychology*, 1969, **81**, 275–280.

SMITH, F. Familiarity of configuration vs. discriminability of features in the visual identification of words. *Psychonomic Science*, 1969, **14**, 261–262.

SMITH, F. *Understanding reading*. New York: Holt, Rinehart & Winston, 1971.

SMITH, F., LOTT, D., & CRONNEL, B. The effect of type size and case alternation on word identification. *American Journal of Psychology*, 1969, **82**, 248–253.

WOODWORTH, R. S. *Experimental Psychology*. New York: Holt, 1938.

REFERENCE NOTES

1. BROOKS, L. *Visual pattern in fluent word identification.* Paper presented at the conference on reading sponsored by Brooklyn College and The Graduate Center of CUNY, 1974.
2. TAYLOR, G. A., MILLER, T. J., & JUOLA, J. F. *Isolating visual units in the perception of words and nonwords.* Paper presented at the annual meeting of the Psychonomic Society, Denver, Colorado, November 1975.
3. ADAMS, M. J. *Perceptual components of the word apprehension effect.* Studies in Attention and Learning, No. 8. Providence, R.I.: Brown University, Walter S. Hunter Laboratory in Psychology, 1976.
4. LABERGE, D. Perceptual learning and attention. To appear in W. K. Estes (Ed.), *Handbook of Learning and Cognitive Processes*, Vol. 4.
5. LANDAUER, T. K., DIDNER, R. S. & FOWLKES, E. B. *Processing stages in word naming: reaction time effects of letter degradation and word frequency.* Paper presented at the 43rd annual meeting of the Eastern Psychological Association, Philadelphia, Pa., 1975.

(Received July 2, 1976)