Journal of Experimental Psychology:
Human Learning and Memory
1981, Vol 7, No. 4, 269-282

0096-1515/81/0704-0269$00.75

Input, Decision, and Response Factors in Picture–Word Interference

Stephen J. Lupker and Albert N. Katz
University of Western Ontario, London, Ontario, Canada

Two variations of the picture–word analogue of the Stroop task were examined in an effort to gain a better understanding of the processes involved in responding to picture–word stimuli. Four stages in this process were outlined and then evaluated as potential sources of the interference in these types of tasks. In Experiment 1 subjects were required to respond yes or no (vocally or manually) to whether the picture was that of a dog. In Experiment 2 subjects were asked to respond by naming the picture's semantic category. Taken together, the results of these experiments indicate that (a) input factors contribute very little to the interference observed, (b) in certain situations some of the interference is due to an interaction of the semantic information from the word and the picture during a decision process, and (c) the response selection and output processes account for most of the interference but only in situations in which the word's name is potentially a response. Implications of these results for the study of automatic semantic processing of words are discussed.

As Quillian (1968) pointed out over 12 years ago, an individual's storehouse of information about any well-known concept is virtually limitless; yet of this information that we all seem to possess, much is clearly irrelevant to our understanding of that concept. That is, in a typical situation where an observer sees an abstract representation of a concept, for example, a picture of a dog in a magazine or the word "dog" in text, the concept "dog" will be understood in the appropriate manner with little irrelevant information ever being retrieved. In addition, the information that is retrieved, that is, the information that allows the concept to be understood, seems to become available essentially automatically with little, if any, effort on the part of the observer.

The processes involved in retrieving information about particular concepts have received much attention in recent years, particularly when words are employed as the means of accessing memory. For example, considerable information has been obtained concerning decisions about whether a letter string is a word or not, whether the concept represented by a word has a particular property or belongs to a particular semantic category, and so on (see Posner, 1978). However, in most of these situations active retrieval strategies are being implemented that are clearly different from whatever retrieval operations are naturally used in reading text. As such, the question of generalizability can be raised. A different way to examine these issues might be to find a task in which some nonintentional processing of a word is inevitable but the required response should not be based on any aspect of the word itself. In such a task, any processing the word receives will be the result of automatic retrieval operations. Thus, any influence the word has on task performance should permit a clearer view of these operations.

In recent years an experimental paradigm has emerged that appears to meet these qualifications, the picture–word interference paradigm. In the basic picture–word interference task, a line drawing ("picture") with a word superimposed is displayed to a subject. The subject's task is to ignore the word and name the picture as rapidly as possible. Typically, it will take a subject 70–90 msec longer to name a picture with a word su-

This research was supported by Grant A633 from the Natural Sciences and Engineering Research Council of Canada. The authors would like to thank Anne Kennedy, Eric Cartman, and Mike Callahan for their help in the data collection and analyses.

Requests for reprints should be sent to Stephen J. Lupker, Department of Psychology, University of Western Ontario, London, Ontario N6A 5C2, Canada.

perimposed than to name the same picture presented alone (Lupker, 1979). This result is, of course, not surprising, since this task seems to be almost totally analogous to Stroop's (1935) color–word interference task. However, what is surprising, as well as more relevant to the present discussion, is the fact that picture naming latency is further prolonged 24–30 msec when the superimposed word names a member of the picture's semantic category (Lupker, 1979; Rosinski, 1977).

This effect, which is referred to as the "semantic category effect," demonstrates nicely that in this task, as in normal reading, certain semantic information becomes available automatically whenever a word is perceived. After all, the subject's basic task is simply to ignore the word, and as such, it would be expected that the word would be processed no more deeply than is absolutely necessary. If this processing were to only the orthographic or phonetic level, there should be no difference between the same semantic category words and unrelated words. The fact that the semantic category effect exists clearly indicates that the word provides a substantial amount of semantic information also.

If the picture–word interference paradigm is to be a useful tool in understanding the automatic retrieval of semantic information, a workable model of the processes involved in the task must be obtained. The present set of experiments represents a further attempt to understand the basic processes involved in responding to picture–word stimuli. The general framework for analyzing this type of task is derived from the standard processing model initially described by Sternberg (1969). In this conceptualization, four major stages or processes can be described: an input process, a decision process, a response selection process, and a response output process. Presumably any of these processes, either singly or in combination with the others, could give rise to the basic interference effect as well as the additional interference observed when the word and the picture represent concepts from the same semantic category.

The first of these processes, the input process, involves the basic perceptual analysis of the display, that is, the resolution of figure and ground. This process is presumed to be precategorical in nature and, in line with the arguments presented by Lupker and Massaro (1979), to be potentially capacity limited (although a strong case to the contrary has been made by Shiffrin, 1975). Interference at this point would be primarily structural because it would be due to this limitation of processing capacity. In the present task, if the picture and the word have to share perceptual resources, the processing of the picture may be prolonged. This type of explanation was initially offered by Hock and Egeth (1970) as a means of explaining the Stroop effect. A more sophisticated version of this explanation could also account for a gradient of interference. Through training or priming, irrelevant components might automatically attract additional processing resources, thus robbing the relevant component of the use of these resources. To the extent that an irrelevant component (e.g., a to-be-ignored word) attracts resources, processing of the relevant component (e.g., a to-be-identified picture) would be further delayed. This notion will be evaluated as a possible explanation of the results of Experiment 1.

In the second process, the decision process, subjects must consider the relevant information provided by the picture in order to make the decision the task demands. If, for example, the subjects are asked to decide whether the picture is that of a dog, as in Experiment 1, they must examine the acquired perceptual information until sufficient evidence, either positive or negative, is accumulated. This description of the decision process is similar to that suggested by McCloskey and Glucksberg (1979). Interference may arise here because the word, in addition to its name, automatically supplies certain semantic information. This information may interact to various degrees with the relevant information from the picture, causing any decision about the picture to be more difficult to make. Unlike the interference discussed above for the input processes, the interference here would not be due to structural limitations but to difficulties in keeping the sources of the information separate. It may be that the interference that arises here would be very task dependent. For different tasks, different types of pictorial information would be relevant to the decision. The degree to which a superimposed word will interfere may be a function of the extent to which its automatically supplied semantic information is relevant to that decision. This general type of explanation was recently invoked by Shaffer and LaBerge (1979) to explain the interference they observed in a somewhat different interference paradigm.

The third process in which interference may arise is response selection. In the response selection process an output code must be formed based on the appropriate decision (e.g., a phonetic representation in a naming task, a motor code representation in a button-pressing task, and so on). Like the decision process, this process should be task dependent. Factors such as the compatibility of the mappings of the response code to a given decision and the number of possible responses may influence the speed with which the subject can complete this process. One can conceptualize interference arising at this point if the word suggests an alternative response. In the standard Stroop or picture–word interference task, the word would, of course, promote its name as a possible response. A gradient of interference may arise if the response suggested by one type of word is more plausible than the response suggested by a second type.

Finally, a fourth process that may be responsible for interference would be the response output process. Interference here is assumed to be due to a structural limitation in the output buffer. According to this notion, not only does the word suggest a legitimate but incorrect response but also the tendency to produce that response is automatically evoked and must be suppressed. For example, in the basic picture–word interference or Stroop task, the word's name may become available before the correct response and may assume a preeminent position in the output buffer. Before the appropriate response can be produced, the tendency to emit this incorrect response must be overcome, leading to a longer response time. To the extent that this incorrect response may be a primed response, overcoming this tendency may be more difficult and more interference may result. Apparently, this type of explanation of the Stroop effect has been suggested previously by numerous investigators (Dyer, 1973; Keele, 1972; Klein, 1964; Posner & Snyder, 1975). However, in reading the accounts offered by these theorists, one cannot be certain whether they are referring to this process or the response selection process as the "output" process. In fact, the two processes seem to be so closely tied that perhaps it has not been deemed important to make the distinction. However, as this analysis hopefully makes clear, these processes are not identical. In fact, it should be possible to create a situation where a superimposed word could suggest an incorrect response, which must be considered by the response selection process, but which would never reach the output level. We suggest that such could be the case in Experiment 1.

Experiment 1

Experiment 1 was designed to investigate the input and decision processes in picture–word interference in a situation where the effects of response selection and response output could be controlled. The subjects were still required to identify the pictured concept, but now, rather than produce the picture's name, they were asked to make a simple yes/no judgment of whether or not the picture was that of a dog.

The subjects were required to respond either by saying yes or no or by pressing one of two buttons. Since only two responses were possible, the only situation in which response output problems could arise would be when the word and the picture were compatible with different responses. However, the superimposed words should not directly evoke the tendency to either say yes or no or push a particular button. Thus, once a response is selected, it should be output without further problems. Potentially, in the vocal task, the tendency to produce the word's name itself could be directly evoked, producing either selection or output problems.

However, due to the fact that only two responses (yes and no) are legitimate and these words are both highly familiar and easily articulable, the word's name should not be considered by the response selection process, nor should it have the opportunity to slip into a preeminent position in the output buffer. This argument is reinforced below by the finding that the observed pattern of interference in Experiment 1 was independent of the method of responding.

In addition to the picture alone condition, there were four interference conditions in Experiment 1. In Condition 1, pronounceable nonwords were superimposed on the pictures. These letter strings should neither supply semantic information that could influence the decision process nor suggest a response of any sort during the response selection process. Thus the only interference they should produce on either positive or negative trials should be during input. In Conditions 2 and 3 the names of nonanimals and animals, respectively, were superimposed on the pictures. These words should cause equivalent problems for the response selection process. That is, they may promote a negative response, which may produce interference on positive trials but not on negative trials. Any differences between these two conditions would then be attributable to either the input or the decision process. If the semantic category effect in the standard picture–word interference task is due to input factors, animal words should cause more interference than nonanimal words on both positive and negative trials. However, if a difference is observed only on positive trials, this would be evidence that the semantic category effect is attributable to decision processes. As suggested above, this effect would be presumed to occur because the semantic information automatically available from the animal words is interacting with the information from the dog pictures to create the decision problems. In Condition 4, the word "dog" was superimposed on the pictures. This word would be expected to produce interference during input on both positive and negative trials to at least the same extent as that produced by the animal words. However, on positive trials this word may facilitate both the decision and the response selection processes, while on negative trials it may inhibit these processes. An analysis of the other interference conditions will allow an evaluation of the contribution of each of these processes to picture–word interference. The results for Condition 4 can then be evaluated from these baselines in order to determine whether facilitation can occur in this task as well as the processes that are affected by the word "dog."

The above remarks, particularly those concerning the decision and input processes, are based on one important assumption: The automatic retrieval of semantic information through words can occur before the subject can make what appears to be a relatively simple judgment, whether or not the picture is that of a dog. To make this decision, subjects do not have to process the picture to the same depth as they do in a naming task. Instead, they should be able to respond based on the presence or absence of a few semantic attributes, which should be readily available from the picture. Certain theories about the relative accessibility of semantic information from words and pictures (e.g., Potter & Faulconer, 1975) seem to imply that it will take longer to obtain semantic information from the word than to derive these few semantic attributes from the picture. If so, no semantic effects should obtain.

Method

Subjects. Forty University of Western Ontario undergraduate volunteers (17 males and 23 females) received course credit for participating in this experiment. Twenty (6 males and 14 females) participated in the vocal version of the task, and 20 (11 males and 9 females) participated in the manual version. All were native English speakers.

Materials and equipment. Twenty line drawings ("pictures") were obtained from children's coloring books. Ten of these were pictures of dogs (the positive trials), 5 were pictures of other four-footed animals, and 5 were pictures of inanimate objects (the negative trials). Five sets of the 20 pictures were produced. Each picture was glued on a 23 × 25.6 cm card. These 5 sets of pictures corresponded to the 5 experimental conditions.

For Condition 1, 10 pronounceable nonwords were created, and each was superimposed on 1 dog picture and 1 nondog picture. For Condition 2, 10 new nonanimal names were selected, and each name was superimposed on 1 of the dog pictures and 1 of the nondog pictures. For Condition 3, 10 new animal names were selected, and each was superimposed on 1 of the dog pictures and 1 of the nondog pictures. For Condition 4, the word "dog" was superimposed on each of the 20 pictures. Condition 5 was the "picture alone" condition, in which no letters appeared on the pictures.

Because the data were available, an effort was made to select words in Conditions 2 and 3 that were equivalent on the dimensions of imageability and printed familiarity using Paivio's (Note 1) norms. The mean imageability ratings for the words in Conditions 2 and 3 were 6.58 and 6.56, respectively. The mean familiarity ratings for the words in Conditions 2 and 3 were 5.04 and 5.12, respectively. (The corresponding ratings for the word "dog" were 6.53 for imageability and 6.33 for familiarity.) In addition, the mean lengths of the words in both conditions and the mean lengths of the pronounceable nonwords were exactly the same. The names of the pictures, the words, and the pronounceable nonwords used are given in the Appendix.

A Ralph-Gerbrands Co. (Model 1-3B-1C) three-field tachistoscope was used to present the stimuli. Viewing distance was 77 cm, and viewing was binocular. The letters typically subtended a visual angle of .24° horizontally and .36° vertically. The pictures subtended visual angles of between 1.90° and 5.74° horizontally and 3.84° and 5.74° vertically. A Hunter-Klockounter (Model 120) timer was used to time subjects' responses. For the vocal responders, an Electro-Voice Inc. (Model 621) microphone was positioned approximately 7 cm from the subject's mouth. The microphone was connected to a Lafayette Instruments Co. (Model 18010) voice-activated relay, which stopped the timer at the initiation of the subject's vocal response. For the manual responders, a board mounted with two telegraph keys was placed on the table in front of the subject. Depression of either of the keys likewise stopped the timer.

Procedure. Subjects were tested individually. All subjects in the vocal condition participated in an unrelated picture–word interference experiment first.[1] Subjects in the manual condition participated in their own experiment twice. Many of these subjects seemed to have trouble getting accustomed to the experimental setup, so the first block was considered a warm-up and only the second set of trials was analyzed. Before the present experiment started, subjects were informed that they would be seeing a series of pictures, some of which would have words superimposed, and their job would be to classify each picture as to whether it was a dog or not. When the picture of a dog was presented they were to respond yes (press the right-hand button with their right index finger), and when any other picture was presented they were to respond no (press the left-hand button with their left index finger). Next, in order to demonstrate what they were being asked to do, they were shown a picture of a cat and told, "For example, if this picture were to appear, your job would be to respond 'no.' " The subjects then responded to each of the 100 stimuli in a random order. For the manual group the stimuli were then shuffled and presented again. Onset of the stimulus started the timer, which was stopped by the subject's vocal or manual response. Each stimulus remained in view for 750 msec, regardless of the subject's reaction time. The interval between stimulus presentations was used by the experimenter to record the response latency and to reset the equipment for the next trial. Thus, this time interval was not held constant but was generally around 5 sec. Errors were recorded, and those pictures were randomly placed back into the set of to-be-presented stimuli. For both groups the entire procedure took about 1 hr.

Table 1
Reaction Times (in msec) as a Function of Experimental Conditions, Experiment 1

Condition	Reaction time		
	Vocal	Manual	*M*
Positive trials			
Word "dog"	604	486	545
Animal name	659	523	591
Nonanimal name	629	506	567
Pronounceable nonword	628	496	563
Picture alone	618	486	552
Negative trials			
Word "dog"	644	537	591
Animal name	627	511	569
Nonanimal name	626	503	565
Pronounceable nonword	637	503	570
Picture alone	616	503	560

Results

As is typically the case in vocal reaction time tasks, errors in the vocal condition were virtually nonexistent (11 errors in all 2,000 trials). The results for these few error trials were not analyzed. There were twice as many errors (22) in the manual condition, and although these errors were generally scattered, a slight pattern emerged inasmuch as 9 occurred in the animal word condition on positive trials. Thus, the reported mean for this condition could be, if anything, slightly deflated.

The mean correct reaction times for the five conditions for both negative and positive trials for both the vocal and the manual responders are presented in Table 1. Each of the individual means is based on 200 observations. For the positive trials, a simple two-way analysis of variance (ANOVA) was performed. The analysis revealed highly significant effects of both Mode of Response, *F*(1,

[1] The task in this experiment was the standard picture naming task. Phonetic, rather than semantic, factors were being investigated, and none of the pictures in the two experiments were the same.

Table 2
Reaction Times (in msec) as a Function of Experimental Conditions, Experiment 1, Negative Trials

	Reaction time	
Condition	Animal pictures	Nonanimal pictures
Word "dog"	625	556
Animal name	598	539
Nonanimal name	586	543
Pronounceable nonword	589	550
Picture alone	583	536
M	596	545

38) = 24.36, $p < .001$, and Word Conditions, $F(4, 152) = 12.97$, $p < .001$, but not a hint of an interaction between these factors ($F < 1.0$).[2] The former effect is due to the fact that subjects were 129 msec faster in the manual condition. The latter effect was investigated more closely through the use of a Newman-Keuls analysis. This analysis revealed that Condition 3, the animal word condition, was significantly different from the other four conditions ($p < .01$). Also, the word "dog" condition was significantly different from the nonanimal word condition ($p < .05$), while the nonanimal word–picture alone and pronounceable nonword–word "dog" comparisons were marginally significant ($.10 > p > .05$).

For the negative trials, a three-way ANOVA was performed with Foil Type (animal vs. nonanimal pictures) as the additional factor. All three main effects—Response Mode, $F(1, 38) = 19.91$, $p < .001$; Foil Type, $F(1, 38) = 217.47$, $p < .001$; and Word Conditions, $F(4, 152) = 7.23$, $p < .001$—were highly significant. As before, the main effect of Response Mode arose because subjects were much faster (119 msec) responding manually than responding vocally. The significant effect of foil type indicated that it was much more difficult to respond negatively to the picture of an animal ($\bar{X} = 596$) than to the picture of a nonanimal ($\bar{X} = 545$). This difference was qualified by a significant Response Mode × Foil Type interaction, $F(1, 38) = 7.96$, $p < .01$, with the subjects in the manual condition showing a larger (61 msec) effect of foil type than that shown by the subjects in the vocal condition (41 msec). The main effect of conditions was again examined more closely through the use of a Newman-Keuls analysis. This analysis revealed that the word "dog" condition differed from the other four conditions ($p < .01$); however, no other differences even approached significance. Neither the Response Mode × Word Conditions interaction, $F(4, 152) = 1.34$, $p > .25$, nor the three-factor interaction, $F(32, 1{,}216) = 1.04$, $p > .40$, approached significance. However, the Word Conditions × Foil Type interaction was marginally significant, $F(4, 152) = 2.19$, $.10 > p > .05$. This interaction, shown in Table 2, indicates a trend for the effect of foil type to be slightly larger in the word "dog" condition (69 msec) than in the other four conditions (39, 43, 47, and 59 msec for the pronounceable nonword, nonanimal word, picture alone, and animal word conditions, respectively).

Discussion

The purpose of Experiment 1 was to evaluate the contributions of the earlier processes to picture–word interference in situations that should be free of response output problems. The technique for eliminating output interference was to use responding procedures quite divorced from any response tendencies the superimposed word could evoke. The success of this manipulation is indicated by the fact that the pattern of interference on both positive and negative trials was independent of response mode (i.e., whether a manual or a vocal response was required).

[2] Because of the arguments presented by Wike and Church (1976), among others, words were not treated as a random factor, as suggested by Clark (1973), in this or any subsequent analyses. In particular, the words used could in no way be thought of as being randomly selected. In Experiment 1 the words in Conditions 2 and 3 were selected in a way to ensure that they were equivalent on the dimensions of imageability and familiarity according to Paivio's (Note 1) norms and to ensure that they were the same length. For Experiment 2, it was again required that the words in the various conditions be equivalent in length. However, because the imageability and familiarity data were not available, Battig and Montague's (1969) production norms were consulted in order to ensure that all the words selected represented highly typical members of their semantic categories.

Three major findings are evident in the present study. The first involves the pronounceable nonword condition. Since these letter strings have no meaning, they should not lead to decision or response selection problems; hence, any interference they do produce can be attributed to the input process. On both positive and negative trials, pronounceable nonwords produced only a nonsignificant 10 msec of interference with respect to the pictures alone. Thus, it appears that general input problems play only a very small role in the processing of picture–word stimuli. The second finding concerns the nonanimal words on positive trials. Any difference between this condition and the pronounceable nonword condition could be attributable to either the decision or the response selection process. The finding was that these conditions did not differ. With respect to the response selection process, these words could presumably have automatically suggested a negative response. The lack of interference here shows, however, that this type of problem did not arise. With respect to the decision process, these words could have failed to cause interference either because the information they automatically supply could be ignored by this process or because it was not available rapidly enough to cause any interference. An examination of the animal word condition, where interference was observed, indicates that the former conclusion is most likely the correct one.

The third, and most important, finding was the significant 24 msec difference between the animal and nonanimal word conditions on positive trials. Since these two conditions should have had equivalent effects at the response selection level, this difference was presumably due to either the input or the decision process. An input explanation would hinge on the idea that something in the nature of the present task, for example, its emphasis on dogs or the extensive use of animal words, could prime those words so that upon recognition they would draw a large amount of attention to themselves and away from picture processing. This argument predicts, however, that this difference should be independent of whether the picture requires a positive or a negative response. Since this difference was only present on positive trials, an explanation of this sort seems unlikely.[3] Instead, it appears that this difference is attributable to the decision process. Apparently the semantic information from the words is available rapidly enough to cause decision problems in the present task. However, problems arise only when this information matches fairly well with the pictorial information the subject is evaluating. In this situation, subjects apparently have difficulty determining which information is relevant to the word and which is relevant to the picture. Thus, decision time is prolonged. When the word and the picture are not semantically similar, as in the nonanimal word condition on positive trials, the two sources of information can be kept separate and no decision problems will be produced. In addition, it should be noted that a second criterion for the appearance of decision interference is that the two sources of information must be compatible with different decisions. In the animal word condition on negative trials with animal foils, the information from the word is also quite similar to the information available from the picture. Yet there is little evidence of interference here, the difference between this condition and the pronounceable nonword condition being only 9 msec.

The results from the word "dog" condition reinforce these conclusions. Although this condition produced little noticeable facilitation on positive trials, it did produce a large amount of interference on negative trials. Further, there was a marginally significant Foil Type × Word Conditions interaction, with the inhibition caused by the word "dog" being much more pronounced for animal as opposed to nonanimal pictures. In fact, for nonanimal pictures this condition differed very little from all of the other conditions. This result also indicates that simply using words incompatible with the correct response was insufficient to produce noticeable interference. It seems that both response incompatibility and semantic simi-

[3] Actually, if a general attending response were being evoked by the animal words, it could interfere at any point in picture processing (see Shiffrin & Schneider, 1977). However, this type of explanation is also discouraged by the fact that the difference was obtained only on positive trials.

larity of picture and word are required for decision problems to arise.

In summary, the following conclusions can be offered. First, input interference played a very small role in the present task, accounting for no more than 10 msec of interference. Thus, by extension, input problems probably also play a fairly small role in the standard picture–word interference task. Second, decision problems did play a role in the present task, but it appeared to be a somewhat restricted role. In this task the semantic information automatically supplied by the words was available rapidly enough to interfere with the subject's decision process. However, interference only occurred when this information was substantially similar to the pictorial information the subject was evaluating and, apparently, when the two sources of information were not compatible with the same decision. Again, by extension, the same argument probably applies to the standard picture–word interference task. The semantic information from words unrelated to the pictures they appear on produces few, if any, problems for the decision process, while words naming members of the picture's semantic category do cause decision problems. Because of the fact that the size of the difference between the animal and nonanimal word conditions on positive trials (24 msec) is essentially the same as the size of the semantic category effect in the standard task, it seems likely that the semantic category effect is due entirely to decision processes. Finally, the response selection process did not appear to cause problems in the present task, the main evidence for this conclusion being the essentially nonexistent difference between the nonanimal word and pronounceable nonword conditions on positive trials.

We should note, however, that the argument is not being made that response selection does not contribute to the interference observed in the standard picture–word interference task. In fact, along with the response output process, selection apparently accounts for the lion's share of the interference. The crucial difference is that in the present task, the permissible responses were quite divorced from any aspects of the to-be-ignored words. In Experiment 1, subjects were permitted only to say yes or no or were required to push one of two buttons. The mapping between, say, a word like "cat" and the word "no" or a left-hand button press is a very artificial one and may require extensive practice to become automatized. In either situation it may not be overly surprising that the superimposed words did not suggest a response. However, in the standard task, the mapping between the word "cat" and the very legitimate response of its name would be a much more compatible one. As such, response selection problems may very well arise when subjects are required to name the pictures. However, at present it is not possible to say how much of the 70–90 msec general interference effect actually is attributable to response selection and how much is attributable to the response output process.

Experiment 2

The results of Experiment 1 suggest two interesting conclusions that deserve further evaluation. The first is that words do not appear to cause problems for the response selection process by suggesting responses other than their own names. This conclusion seems to be at odds with the theorizing of past investigators (Keele, 1972; Warren, 1972). In particular, Warren has shown that the prior presentation of a short list of category members can at least prime the category's name, so if it subsequently appears as a Stroop stimulus additional interference is observed. Thus, Warren's results suggest that at least in some contexts words may automatically suggest their category's name. This possibility was investigated in Experiment 2. The task which was used was a categorization task in which the subject was required to produce the name of the picture's semantic category. A set of five categories was defined as being "active," that is, as being the categories from which all of the pictures were selected. Thus, these category names were the only legitimate responses. This manipulation should have the effect of priming these five words so that if they appear as the superimposed words they will cause more problems for the response selection process than the names of "inactive" categories (Posner & Snyder, 1975; Warren, 1972). The question is whether this same difference will obtain between the names of members of the active categories and the names of the members of the inactive categories.

The second conclusion from Experiment 1 to be considered further is that the automatically available semantic information from the word only causes problems in making decisions (a) when the word can supply information similar to but obviously not identical with that available from the picture and (b) when the two stimulus components are not compatible with the same decision. The task used in Experiment 2 allowed an examination of the first of these criteria in a somewhat different context. That is, a condition was created in which the word and the picture represented members of the same semantic category. In this condition the information provided by the two sources should be highly similar; however, the two components will be compatible with the same decision. The question is whether satisfying only the one criterion will produce problems for the decision process.

Experiment 2 entailed the use of six conditions. In Condition 1, the word was the name of an incongruent active category. In Condition 2, the word was the name of an inactive semantic category. In Conditions 3 and 4, the names of members of incongruent active and inactive categories, respectively, were used. In Condition 5, the word and picture represented members of the same semantic category. Finally, there was a picture alone control condition. In addition, in an effort to further facilitate comparisons a seventh condition was added. The logic for this condition is as follows. Recall that in Experiment 1 response selection and output problems were controlled by using in the one instance button-press responses and in the other instance a small set of very familiar vocal responses. In Experiment 2, the response set has been increased to five responses, none of which are extremely familiar. As such, there are now potential response output problems in the category members conditions caused by the word's name itself (cf. Williams, 1977). That is, the word may suggest itself as an appropriate response at the response selection stage or create output problems by slipping into the output buffer. In order to provide an alternative control, this seventh condition was added in which the superimposed word was the picture's name. This identity condition should provide the same potential for output interference as the other category members conditions. Thus, it appears to be a better baseline against which to compare the other conditions.

Method

Subjects. Eighteen University of Western Ontario undergraduate volunteers (7 males and 11 females) received course credit for participating in this experiment. All were native English speakers.

Materials and equipment. Battig and Montague's (1969) production norms were consulted in order to find five concrete semantic categories, each having four easily picturable members among its 15 most frequently produced instances. The categories selected were animals, clothing, fruit, furniture, and vehicles. Line drawings ("pictures") of each of the four instances of each category were selected from children's coloring books. Seven sets of these pictures were produced, and each picture was glued on a 23 × 25.6 cm card. These seven sets of pictures corresponded to the seven experimental conditions.

To create the identity condition, the picture's name was superimposed on each picture. To create the same semantic category members condition, Battig and Montague's (1969) norms were again consulted to find four other instances from among the 15 most frequent from each category. Each word was superimposed on one of the pictures from the same semantic category. To create the active category members condition, Battig and Montague's norms were once again used to find four other instances from among the 15 most frequent from each category. From each of these sets of four, one word was superimposed on one picture from each of the other four categories. To create the active category condition each category name was superimposed on one picture from each of the other categories. To create the inactive category condition five other relatively concrete categories were selected (buildings, flowers, tools, vegetables, and weapons) and each of these category names was superimposed on four pictures, all from different categories. To create the inactive category members condition, four instances of each of the categories used in the previous condition (except buildings) were selected from the 15 most frequent instances in Battig and Montague's norms. Four common buildings were also selected. Each of these words was superimposed on one of the pictures, with the restriction that no two words from the same semantic category were to be on pictures that belonged to the same semantic category. Finally, in the picture alone condition, no words appeared on the pictures.

Because the data were not available, it was not possible to equate the words in the various conditions on the dimensions of imageability and printed familiarity.

However, because of the fact that only frequent instances from Battig and Montague's norms were used, it would have been surprising if the words differed greatly on either of these dimensions. An estimate of the familiarity of the words was obtained by consulting Kucera and Francis' (1967) word frequency norms. The median frequencies for the identity, same semantic category members, active category members, and inactive category members conditons were 18, 23, 13, and 34, respectively. The category names were, perhaps, a bit more frequent with the active category names having a median frequency of 34 and the inactive category names having a median frequency of 40. In addition, the mean word lengths in the identity, same semantic category members, active category members and inactive category members conditions were essentially equal, ranging from 4.6 to 5.2 letters. The mean lengths of the category names were slightly longer, 7.0 letters in the active category condition and 6.6 letters in the inactive category condition. The names of the pictures and the words used are given in the Appendix.

The tachistoscope, timer, microphone, and voice-activated relay were the same as those used in Experiment 1. Also, the letters and pictures were the same size as those used in Experiment 1.

Procedure. The procedure was similar to that used in Experiment 1. Before the experiment started subjects were informed that they would be seeing a series of pictures, most of which would have words superimposed, and their job would be to produce the picture's category name. They were told that the only categories used in the present experiment were animals, clothing, fruit, furniture and vehicles and that they should restrict their responses to these five category names. They were encouraged to respond as rapidly and accurately as possible. They were then shown the picture of a dog and told, "For example, if this picture were presented, your job would be to respond 'animal.'" The subjects then responded to each of the 140 stimuli in a random order. As before, errors were recorded, and those pictures were randomly placed back into the set of to-be-presented items. Following a brief (3 min.) rest, the subjects once again responded to each of the 140 stimuli in a different random order. The entire procedure took about 1 hr.

Results

Once again, errors were basically nonexistent (35 in 5,040 trials, and no more than 6 in any condition). Thus, the results from the few error trials were not analyzed. The mean correct reaction times for all seven conditions are presented in Table 3. Each of these means is based on 720 observations.

A three-factor ANOVA (Conditions × Trial Block × Response) was performed on the data. As is obvious, the main effect of conditions was highly significant, $F(6, 102) = 6.31, p < .001$. A subsequent Newman-Keuls analysis revealed that the mean for the picture alone condition was significantly less than the means for all other conditions ($p < .05$) and that the mean for the active category condition was marginally greater than the means for all other conditions ($p < .06$). No other differences approached significance. Also significant was the main effect of trial block, $F(1, 17) = 4.66, p < .001$, indicating that subjects improved with practice. However, this factor did not interact with conditions, $F(6, 102) = 1.17$, *ns*.

Table 3
Reaction Times (in msec) as a Function of Experimental Conditions, Experiment 2

Condition	Reaction time
Active category name	874
Inactive category name	854
Active category member's name	856
Inactive category member's name	855
Same semantic category member's name	845
Picture's name	843
Picture alone	821

Finally, the main effect of response was highly significant, $F(4, 68) = 26.42, p < .001$ as was the interaction of this factor with conditions, $F(24, 408) = 2.09, p < .01$. The main effect of response was not surprising. Some category names were undoubtedly easier to articulate than others, and the pictures used in the present experiment probably allowed differential accessibility of the various category names. A closer examination of the Response × Conditions interaction indicated that the basic relationship between conditions was essentially replicated for each response, except for the response "animal." However, since there seems to be no reason why this response should be any different from the other four, in particular it was neither the easiest nor the most difficult to produce, and given the statistical power involved in testing this particular interaction, its significance does not seem to be a cause for concern.

Discussion

The first thing to note about Experiment 2 is that interference was observed in the identity condition. Although input problems may account for a small portion of the difference, most of it can be attributed to the word's name competing with the correct response at either the response selection or the output level. Apparently, as hypothesized, the amount of interference attributable to these processes is a function of the uncertainty the processes face. In the present task, response uncertainty was increased from that in Experiment 1 by increasing the set of legitimate responses from two to five. In addition, this response set was much more arbitrarily defined, as it potentially could have included a number of other responses. Thus, the opportunity for the word's name to compete at the selection or output levels was increased and additional interference was observed. In comparison, in the standard picture–word interference task, almost any imageable object can appear as the picture. Thus, response selection and output problems would be expected to be even greater.

The first conclusion from Experiment 1 being reexamined in Experiment 2 was that words do not cause problems for the response selection process by suggesting responses other than their names. In particular, in Experiment 2 the question was whether they would automatically suggest the name of their semantic category. In this task the active categories were "primed" by instructing the subjects to use only those names in making their responses. As expected, the active category names did produce 20 msec more interference than did the inactive category names which were not primed. If the names of active and inactive category *members* suggested their categories' names to the response selection process, a difference should also have been obtained between the two category members' conditions. Clearly it was not. Thus, even though the generation of category names was central to the present task, this same process did not seem to be engaged by the superimposed words.

This finding permits a much clearer evaluation of the other conclusion from Experiment 1 that was also investigated in Experiment 2, namely, that decision problems only arise when (a) the semantic information from the word is similar to but not identical with the semantic information from the picture and (b) the two components are not compatible with the same decision. The first criterion can be examined in isolation by considering the same semantic category members condition. While this condition did produce interference with respect to the picture alone condition, its mean reaction time was virtually identical to that in the identity condition. Therefore, it appears that simply having two components that supply similar but not identical information is not sufficient to produce decision problems. In order to examine the second criterion in isolation, one can compare both the active and inactive category members conditions with the identity condition. The approximately 12 msec difference between these conditions and the identity condition fell far short of significance. As such, although there may have been a small effect of decision processing here, the statistical equivalence among these three conditions argues that satisfying the second criterion alone, that is, having two components supplying information incompatible with the same decision, is also not sufficient to produce decision problems. Therefore, it appears that in the picture–word interference context both of the above criteria must be satisfied before decision problems will emerge.

Conventional wisdom (Potter & Faulconer, 1975) has it that pictures supply semantic category information faster than words do. In fact, Smith and Magee's (1980) recent argument for this conclusion was based, in part, on their results in a similar picture–word interference task. Thus, an alternative interpretation of the lack of decision problems in Experiment 2 would simply be that the words did not automatically supply interfering semantic information before a category decision could be made about the picture. However, the results of Experiment 1 would indicate that an explanation of this sort is unlikely. Reaction times in Experiment 1 were much shorter than those in Experiment 2, indicating a much shorter and more rapid decision process. Nonetheless, semantic interference was observed in Experiment 1. Apparently, automatic semantic processing of words was rapid enough to produce interfering semantic information before subjects could make even this very simple decision about a picture. As such, at least the same amount of potentially interfering semantic information must have been available to the decision process in Experiment 2. However, since both criteria for decision problems were not met, this information produced no interference.

Conclusions

The present experiments were designed in an effort to increase our understanding of the processes involved in responding to picture–word stimuli. Four stages in this process were outlined—input, decision, response selection and response output—and the first three were examined with the emphasis placed on the decision process. The results from the present experiments seem to suggest the following conclusions.

First, the interference in picture–word tasks caused by an input process characterized by perceptual resolution of figure and ground is, at best, minimal. Second, automatic semantic processing of a word will produce interference with the decision process. This interference is somewhat specific, however, and appears only to occur when information supplied by the word is (a) similar but not identical to that supplied by the picture and (b) not compatible with the same decision as being made about the picture. Both of these criteria must be met for interference to occur. The semantic category effect observed in both the standard naming task and the yes/no task in Experiment 1 appears to be due to interference at this level. Third, interference with the response selection process is possible, but only when it is based on the name of the superimposed word and not on other information automatically supplied by the word. In fact, as shown in Experiment 2, superimposed words do not suggest their category names, even in situations where the process of generating category names is continually engaged. Finally, from the present data it is not possible to determine how much interference in the standard task is due to response output. However, given the contributions of the input and decision processes, it seems likely that the two response processes must account for the lion's share of this interference.

Perhaps it is appropriate at this point to ask what has been learned about automatic semantic processing of words. It is clearly a very rapid process. Indisputably, certain semantic information is much more difficult to retrieve when memory is accessed by means of a word than when it is accessed by means of a picture (Potter & Faulconer, 1975). However, the semantic effects in Experiment 1 imply that whatever information the word automatically supplies arrives at least as rapidly as the very basic information about the picture which is necessary to complete this simple task. In addition, the information that becomes available seems to be relatively immutable. The experimental context in Experiment 2 could very easily have led to the retrieval of the words' category names if they had been at all available. Clearly, however, these names were not being retrieved. Thus, the ability of experimental context to alter the kind of information that automatically becomes available from words appears to be somewhat limited. As such, whatever knowledge about automatic retrieval operations can be derived from the picture–word interference paradigm will probably be fairly generalizable to normal reading situations. Finally, whatever semantic information is obtained from a word can generally be kept separate from the semantic information being derived from a picture. The nonanimal words in Experiment 1 undoubtedly supplied the same kind of information that the animal words supplied on positive trials, and the information from both types of words was incompatible with a positive decision. Yet the unrelated words interfered no more than the pronounceable nonwords, which carry no semantic information. Also, in Experiment 2, active category members that should have provided extremely problematic semantic information led to essentially no more interference than words providing entirely compatible semantic information. It appears that unless the semantic information from the word and the picture can interact in some way, the semantic information the word supplies seems to be basically irrelevant to the subject's decision processing. Thus, the effects of automatically retrieved semantic information will not be evident in every task involving picture–word stimuli. Unfortunately, the implication of this final conclusion is that the future usefulness of the picture–word interference paradigm for studying automatic semantic processing of words may be somewhat limited. Nonetheless, the nature of these interactions, whenever they arise, should provide useful insights not only into automatic semantic processing of words but also into the nature of the memory system in which those interactions take place.

Reference Note

1. Paivio, A. *Imagery and familiarity ratings for 2,448 words: Unpublished norms.* Department of Psychology, University of Western Ontario, London, Ontario, Canada, 1971.

References

Battig, W. F., & Montague, W. E. Category norms for verbal items in 56 categories: A replication and extension of the Connecticut category norms. *Journal of Experimental Psychology Monograph*, 1969, *80*(3, Pt. 2).

Clark, H. H. The language-as-fixed-effect fallacy: A critique of language statistics in psychological research. *Journal of Verbal Learning and Verbal Behavior*, 1973, *12*, 335–359.

Dyer, F. H. The Stroop phenomenon and its use in the study of perceptual, cognitive and response processes. *Memory & Cognition*, 1973, *1*, 106–120.

Hock, H. S., & Egeth, H. E. Verbal interference with encoding in a perceptual classification task. *Journal of Experimental Psychology*, 1970, *83*, 299–303.

Keele, S. W. Attention demands of memory retrieval. *Journal of Experimental Psychology*, 1972, *93*, 245–248.

Klein, G. S. Semantic power measured through the interference of words with color-naming. *American Journal of Psychology*, 1964, *77*, 576–588.

Kucera, H., & Francis, W. N. *Computational analysis of present-day American English.* Providence, R.I.: Brown University Press, 1967.

Lupker, S. J. The semantic nature of response competition in the picture–word interference task. *Memory & Cognition*, 1979, *7*, 485–495.

Lupker, S. J., & Massaro, D. W. Selective perception without confounding contributions of decision and memory. *Perception & Psychophysics*, 1979, *25*, 60–69.

McCloskey, M., & Glucksberg, S. Decision processes in verifying category membership statements: Implications for models of semantic memory. *Cognitive Psychology*, 1979, *11*, 1–37.

Posner, M. I. *Chronometric explorations of mind.* Hillsdale, N.J.: Erlbaum, 1978.

Posner, M. I., & Snyder, C. R. R. Attention and cognitive control. In R. L. Solso (Ed.), *Information processing and cognition: The Loyola Symposium.* Hillsdale, N.J.: Erlbaum, 1975.

Potter, M. C., & Faulconer, B. A. Time to understand pictures and words. *Nature*, 1975, *253*, 437–438.

Quillian, M. R. Semantic memory. In M. Minsky (Ed.), *Semantic information processing.* Cambridge, Mass. M.I.T. Press, 1968.

Rosinski, R. R. Picture–word interference is semantically based. *Child Development*, 1977, *48*, 643–647.

Shaffer, W. O., & LaBerge, D. Automatic semantic processing of unattended words. *Journal of Verbal Learning and Verbal Behavior*, 1979, *18*, 413–426.

Shiffrin, R. M. The locus and role of attention in memory systems. In P. M. A. Rabbitt & S. Dornic (Eds.), *Attention and performance V.* London: Academic Press, 1975.

Shiffrin, R. M., & Schneider, W. Controlled and automatic human information processing: II. Perceptual learning, automatic attending and a general theory. *Psychological Review*, 1977, *84*, 127–190.

Smith, M. C., & Magee, L. E. Tracing the time course of picture–word processing. *Journal of Experimental Psychology: General*, 1980, *109*, 373–392.

Sternberg, S. The discovery of processing stages: Extentions of Donders' method. *Acta Psychologica*, 1969, *30*, 276–315.

Stroop, J. R. Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, 1935, *18*, 643–662.

Warren, R. E. Stimulus encoding and memory. *Journal of Experimental Psychology*, 1972, *94*, 90–100.

Wike, E. L. & Church, J. D. Comments on Clark's "The language-as-fixed-effect fallacy." *Journal of Verbal Learning and Verbal Behavior*, 1976, *15*, 249–255.

Williams, E. The effect of amount of information in the Stroop color word test. *Perception & Psychophysics*, 1977, *22*, 463–470.

Appendix follows on next page

Appendix

Table A1
Stimuli in Experiment 1: Positive Trials

Picture	Word: Dog	Word: Animal	Word: Nonanimal	Pronounceable nonword
Dog	DOG	CAT	DAWN	LURIM
Dog	DOG	HORSE	FLASK	ZESAM
Dog	DOG	LION	GARDEN	GADID
Dog	DOG	ELEPHANT	GEM	SYDAH
Dog	DOG	BEAR	HAMMER	COKEM
Dog	DOG	FOX	LEMON	PATEK
Dog	DOG	LEOPARD	MAST	MUJIK
Dog	DOG	MULE	OCEAN	VOMER
Dog	DOG	CAMEL	TOWER	DERAY
Dog	DOG	BEAVER	PALACE	PAROG

Table A2
Stimuli in Experiment 1: Negative Trials

Picture	Word: Dog	Word: Animal	Word: Nonanimal	Pronounceable nonword
Mouse	DOG	CAT	FLASK	GADID
Rabbit	DOG	LION	GEM	ZESAM
Pig	DOG	BEAR	LEMON	LURIM
Cow	DOG	LEOPARD	OCEAN	SYDAH
Goat	DOG	CAMEL	TOWER	COKEM
Bed	DOG	HORSE	DAWN	PATEK
Door	DOG	ELEPHANT	GARDEN	MUJIK
Leaf	DOG	FOX	HAMMER	VOMER
Foot	DOG	MULE	MAST	DERAY
Car	DOG	BEAVER	PALACE	PAROG

Table A3
Stimuli in Experiment 2

Picture	IC	ICM	AC	ACM	ID	SSCM
Car	WEAPON	KNIFE	FURNITURE	PLUM	CAR	TRAIN
Plane	TOOL	LIBRARY	ANIMAL	SOCKS	PLANE	BUS
Bike	VEGETABLE	BEAN	CLOTHING	MOUSE	BIKE	SCOOTER
Motorcycle	FLOWER	ORCHID	FRUIT	BUREAU	MOTORCYCLE	BOAT
Bed	BUILDING	GUN	VEHICLE	LIME	BED	DESK
Chair	WEAPON	HAMMER	ANIMAL	JACKET	CHAIR	SOFA
Lamp	TOOL	CHURCH	CLOTHING	RAT	LAMP	RUG
Table	VEGETABLE	VIOLET	FRUIT	WAGON	TABLE	CABINET
Cat	BUILDING	RIFLE	VEHICLE	GRAPE	CAT	SHEEP
Lion	FLOWER	SAW	FURNITURE	SKIRT	LION	COW
Horse	WEAPON	CARROT	CLOTHING	STOOL	HORSE	BEAR
Pig	TOOL	SCHOOL	FRUIT	TRACTOR	PIG	DOG
Hat	BUILDING	SWORD	VEHICLE	LEMON	HAT	SHIRT
Pants	VEGETABLE	DRILL	FURNITURE	TIGER	PANTS	BLOUSE
Shoe	FLOWER	PEA	ANIMAL	COUCH	SHOE	TIE
Dress	WEAPON	ROSE	FRUIT	TRUCK	DRESS	SWEATER
Apple	BUILDING	HOUSE	VEHICLE	COAT	APPLE	PEACH
Banana	TOOL	WRENCH	FURNITURE	ELEPHANT	BANANA	ORANGE
Pear	VEGETABLE	CORN	ANIMAL	DRESSER	PEAR	APRICOT
Cherry	FLOWER	TULIP	CLOTHING	SHIP	CHERRY	PINEAPPLE

Note. IC = inactive category name; ICM = inactive category member's name; AC = active category name; ACM = active category member's name; ID = picture's name; and SSCM = same semantic category member's name.

Received December 8, 1980 ■