FACULTY WORKING PAPER 92-0124

An Examination of the Categorization Process for Pictures and Words

Madhubalan Viswanathan
Department of Business Administration
University of Illinois

Terry L. Childers
Department of Marketing & Logistics Management
University of Minnesota

Bureau of Economic and Business Research
College of Commerce and Business Administration
University of Illinois at Urbana-Champaign

BEBR

FACULTY WORKING PAPER NO. 92-0124

College of Commerce and Business Administration

University of Illinois at Urbana-Champaign

April 1992

An Examination of the Categorization Process for Pictures and Words

Madhubalan Viswanathan is an Assistant Professor of Marketing at the Department of Business Administration, University of Illinois, Champaign, Illinois 61820. Terry L. Childers is an Associate Professor of Marketing at the Carlson School of Management, University of Minnesota, Minneapolis, Minnesota 55455.

The authors would like to thank Professor Joan Gay Snodgrass, New York University and Deborah Roedder John, University of Minnesota for comments on an earlier draft of this manuscript and to thank Paul Bottom, University of Minnesota for assistance in data collection.

Abstract

Categorization processes are fundamental to such critical consumer judgments as identification and brand evaluation. However, little is known about the processes that underlie these judgments. Related to these judgments is the nature of marketing communications that provide information upon which product categorization takes place. An interplay between the categorization process and the modality in which information is provided is proposed in this paper. Differences in the categorization process for pictures versus words are investigated through three laboratory experiments. Results are supportive of a model of categorization which maintains that pictures have faster access than words to their categories and promote faster exemplar to exemplar associations. The proposed model of this process asserts that pictures achieve this advantage through their faster access of the category concept and due to their simultaneous activation of overlapping features between category instances. Results are discussed in terms of their implications for product categorization, effects on decision framing and the configurality of cues, and on the assessment of individual differences in the processing of visual versus verbal information.

Recently there has been an increasing interest among marketers in the area of categorization (Cohen & Basu 1987; Sujan 1985). Categorization appears to play a role in the manner in which people judge product category membership (Cohen & Basu 1987). This in turn translates into implications for both identification and evaluation of a particular brand (Sujan 1985; Cohen, 1981). Cohen & Basu (1987) point out that, while classification procedures are useful in assessing the perceived similarity of products, they provide little insight into the underlying process of categorization. Research aimed at understanding the process of categorization would, thus, be useful to understanding consumer behavior.

A second area of importance to consumer researchers is the modality in which information is presented in marketing communications. Childers & Houston (1984) point out that relatively less attention has been paid to nonverbal information processing and particularly to comparisons between verbal and pictorial information processing. They suggest that consumer research should focus on picture–word effects on consumer memory and judgment. Given that categorization is a pervasive phenomenon which impacts both product identification and product evaluation, the processes involved in picture and word categorization are vital to our understanding of these related areas. In fact it could be stated that an understanding of the categorization process is often a prerequisite to researching picture–word effects since a fundamental difference in the processing of pictures and words may lie in the process(es) by which they are categorized. This argument appears plausible in light of numerous differences that have been found between pictures and words in a number of tasks (Snodgrass 1980). Categorization is one such task of particular importance to consumer researchers.

The purpose of this paper is to report the results of a series of experiments conducted to evaluate alternative views of picture–word categorization. First of all, past empirical research that traces processing differences for pictures versus words is reviewed. Different views of picture and word categorization are presented in the next section and hypotheses are derived for each of these views. This is followed by a description and a discussion of the results of each experiment. Finally, a concluding discussion is provided which interprets the pattern of results obtained in terms of a process model of picture–word categorization.

REVIEW OF EMPIRICAL STUDIES ON PICTURE–WORD PROCESSING

This section provides a review of past research on picture–word categorization. Studies on picture–word categorization can be classified on the basis of their results. Some studies have shown a picture advantage in categorization and these studies are discussed in the first part of this review. Other studies have shown no differences between pictures and words in categorization and these studies are reviewed subsequently.

One of the first studies in the area of picture–word categorization was performed by Potter & Faulconer (1975) who studied the speed of categorization of pictures and words. Here, subjects were exposed to stimuli from eighteen categories in pictorial or verbal form for 250 ms each. They were informed of a category name just before exposure to each stimulus and were required to respond (with a Yes/No) to whether the stimulus belonged to the named category or not. The authors found that pictures were categorized faster than words by an overall difference of 51 ms.

Pellegrino et al. (1977) used a same/different categorical judgment task to evaluate models of semantic memory. Subjects were exposed to a pair of stimuli simultaneously and required to make a judgment whether the pair belonged to the same category or not. Pairs of stimuli were chosen from two categories and consisted of different possible combinations of pictures and words (i.e., picture–picture (referred to as pp), picture–word (pw), word–picture (wp), and word–word (ww)). These combinations refer to left–right ordering during the simultaneous presentation. The patterns of reaction time results were as follows: picture–picture (pp) < picture–word (pw) < word–picture (wp) < word–word (ww). A follow up experiment involved the presentation of single pictures or words and the task was to decide whether these pictures/words were animals or objects. Pictures yielded faster reaction times than words by about 70 ms. The authors concluded that the data conforms to a model where both size and category information is stored in a single memory system, that system being verbal, nonverbal or amodal.

Research on picture–word categorization has not consistently shown a picture advantage. Smith & Magee (1980) used an interference task which required subjects to make a judgment as to whether a picture or a word belonged to a previously named category or not. The reaction time for categorizing pictures versus words was not significantly different. A subsequent experiment also tested recognition memory following a categorization judgment task. Again no significant difference was obtained in both category judgment and recognition memory. It is thus possible that picture–word categorization differences are sensitive to the task required within the experiment.

Snodgrass and McCullough (1986) offer an explanation for a picture advantage in categorization. The authors exposed subjects to pictures or words requiring a judgment as to whether the stimuli belonged to one of two previously named categories. They showed that visual similarity and stimulus modality interact such that a picture advantage in categorization is obtained only when visual similarity is a reliable cue (i.e., high within–category visual similarity and low between–category visual similarity among stimuli used in the experimental task) to category membership. Otherwise, pictures showed a large disadvantage in categorization. In a subsequent experiment, subjects were exposed to pairs of stimuli of pictorial or verbal form shown 1 sec. apart and were required to make a category judgment whether a pair belonged to the same category or not. The results demonstrated a consistent picture advantage for "same" and "different" judgments only when visual similarity was a reliable category cue.

Thus, one possible explanation for a picture advantage in certain categorization tasks could be due to the use of visual similarity in making a decision rather than access to the meaning of an exemplar (Snodgrass & McCullough 1986). Therefore, any *simultaneous comparison* of a pair of pictorial stimuli is open to explanation on the basis of a visual similarity. This is because a simultaneous presentation of a pair of stimuli does not allow an interval between presentation of stimuli (i.e., an inter–stimulus interval) for categorization of a stimulus. Therefore, Ss could perform the experimental task on the basis of visual similarity without categorizing the stimuli. Hence, it is important to design an experimental task which minimizes the use of visual similarity in order to study the categorization process. Pellegrino et al. (1977) generated hypotheses based on the assumption that processing occurs from the left to the right of a screen on which a pair of stimuli are simultaneously presented. Such an approach seems inappropriate given the finding (Snodgrass & McCullough 1986) that visual similarity is used as a cue (even when an interstimulus interval is used). Past research has also required the performance of a task where Ss are primed with a category label and asked to make a Yes/No decision (Smith and Magee 1980; Potter and Faulconer 1975). This provision of a category label as a part of the experimental procedure is also open to alternate explanations (cf., Smith & Magee 1980; Potter and Faulconer 1975). It could be argued that Ss verbalize the category label while performing the categorization task thereby causing interference with word recognition (or the naming of words) but not with picture recognition (Snodgrass 1980).

This review of past research shows that a picture advantage is not consistently obtained in categorization. Further, studies in the past are open to alternate explanations due to the experimental procedures that have been employed. The nature of the experimental procedure appears to be crucial to isolating the task from underlying differences in picture–word categorization. It appears from a synthesis of the above studies, that an "ideal" approach would be to (i) expose Ss to a stimulus, (ii) allow activation of a category, (iii) expose a second stimulus which requires a Yes/No decision, and (iv) prevent the priming of one or a few categories by randomly using a range of categories *without* pre–specification.

The implementation of such an approach, however, poses a few issues in terms of ensuring that Ss do, indeed, follow the "desired" process. For instance, Ss could wait for the subsequent stimulus (category exemplar) and make a decision with or without activating a category. This could be done on the basis of visual similarity rather than categorization. Appropriate instructions are required to ensure that such a process does not take place. Subjects may also activate a category and verbalize it thereby causing interference in word recognition. The inter–stimulus interval (elapsed time between two category exemplars) must be provided such that it is large enough to allow for activation of a category, while discouraging comparison of perceptual features, or categorical comparison after exposure to the subsequent stimulus. However, it should not be so large as to allow time for verbalization of category labels. An experimental paradigm is used here which attempts to incorporate these desired elements.

ALTERNATE CONCEPTUAL VIEWS OF PICTURE–WORD CATEGORIZATION

Several alternate explanations of the categorization process are presented and then hypotheses are generated in the context of the experimental paradigm discussed earlier. Research in the area of picture–word categorization has usually been conducted with a view toward providing evidence on the nature of semantic memory (cf., Pellegrino et al. 1977). Such an approach has led to a focus on the *structure* of semantic memory (i.e., dual versus amodal coding views) while the *process* of picture–word categorization has been neglected. Structural explanations refer to explanations on the basis of coding systems, such as, a dual–code model

Paivio (1971). At the outset, it should be mentioned that the purpose here is not to test different theories of semantic memory in the substantive context of categorization (as was the case in the past) but, *to test different versions of the categorization process*. In other words, the primary focus here is on the categorization process. In doing so, the emphasis here is on processual rather than structural explanations due to the neglect of this facet of research in the past.

A model of semantic memory is required in order to provide a framework for theorizing about picture–word categorization. In this paper, the hypotheses are derived within the framework of Snodgrass's (1980) model of semantic memory. The choice of this model was based on its interplay between process and structure and within the latter, upon its incorporation of different perspectives of semantic memory such as the propositional view (Anderson & Bower 1973) and the dual coding view (Paivio 1971). This model postulates three levels of processing. The first level of processing (Level I) contains raw verbal or visual codes as a result of processing physical characteristics. Level II contains acoustic and visual image stores with prototypical information. Level III contains a propositional store consisting of nodes and interconnections of an abstract nature. The model postulates that acoustic images could access nodes which are not accessible to visual images. This model incorporates the dual coding view with its two image stores in level II and the propositional view with the abstract nature of the level III store. Therefore, this model provides a relatively flexible basis for hypothesis generation. The notion that there is only partial overlap between the propositional space accessed by visual and acoustic images is not very clearly supported in this model. Hence, it will be assumed here that visual and acoustic images have complete access to the propositional store. This assumption, although not critical to our subsequent results, is in keeping with our emphasis on processual rather than structural explanations of the phenomenon of interest.

Each explanation for picture–word processing differences in categorization presented here will be assessed in its pure form without allowing for additional qualifications which might be employed to explain certain isolated phenomenon. Such an approach will facilitate the use of parsimony as a criterion in evaluating various theoretical positions. Qualifications, however, will be addressed when discussing the results. Also, no view will be excluded on qualitative grounds, but will be assessed on empirical grounds. In the following sections three views of picture-word categorization will be considered, and are referred to as **amodal**, **picture-based**, and **verbally-based** categorization.

Amodal View of Picture–Word Categorization

One model of categorization would suggest that the process of categorization is insensitive to the modality of an input (referred to as the “amodal” view). Such a view would imply no differential category access for pictures or words. This view suggests that the process of categorization is equally receptive to pictorial or verbal input. Hence, a picture advantage in categorization would be attributed to encoding operations *prior* to the categorization process (i.e., before access to Level II of the Snodgrass model). Therefore, this view is similar to the arguments presented by proponents of a common coding model to explain a picture advantage in categorization (Pylyshyn 1973;1981). Considering the experimental paradigm detailed earlier, the main hypothesis generated by such a view is that the category decision is independent of the modality of the stimulus. Predictions relating to this view are summarized in Fig. A and Table 1. The thrust of such a view is that the abstract propositional store is equally accessible to pictures and words and that both pictures and words access identical nodes in semantic memory during categorization.

Insert Figure A and Table 1 about here

Picture–Based Categorization

Another view of the categorization process is that it is more receptive to pictorial than to verbal input. It is possible that a pictorial input has to undergo fewer transformations than a verbal input before being subject to the categorization process (referred to in the rest of the paper as the "picture" view). The main argument supporting this view is that a pictorial stimulus is categorized faster than a verbal stimulus. This view would explain previous findings of a picture advantage in categorization in terms of differential transformations for pictures and words during categorization. The specific hypotheses generated by this view are shown in Fig. A and Table 1. The main rationale for these predictions is that, inasmuch as additional transformation is required for words (represented by transfer time "t"), additional reaction times for the categorization of words will be registered. This view suggests that pictures are superior to words by possessing a format more compatible for access into the propositional store for categorization. Both access *into* semantic memory and the nodes accessed *during* categorization could be different between pictures and words.

An issue of importance here relates to the assumptions made about picture and word access to imagistic and acoustic stores, respectively. The assumptions would affect predictions derived for different views of the categorization process. A picture advantage in recognition has been hypothesized to explain a picture advantage in semantic access (Friedman & Bourne 1976). This view has been contested by dual coding theorists in the context of the debate on the nature of semantic memory (te Linde 1982). It appears that any stance on this issue would amount to ruling out consideration of an alternate explanation of picture–word categorization (for example, an assumption of faster picture recognition before semantic access would support explanations for picture–word categorization presented by abstract coding theorists). It is assumed here, for the purpose of deriving predictions, that pictures and words have equal access to the imagistic and acoustic store. Alternate explanations derived by altering this assumption are discussed in the results.

Verbally–Based Categorization

Another possible view of categorization is that the underlying process is more receptive to verbal input (referred to as the "verbal" view). This view would require that pictorial stimuli undergo greater transformations than verbal stimuli before being subject to the categorization process. This view would have to explain a picture advantage in categorization in terms of operations occurring prior to categorization which outweigh a subsequent advantage to verbal input. Such a view would lead to the set of predictions shown in Fig. A and Table 1. These predictions are derived using similar arguments to those presented for the "picture" view.

Augmenting Categorization with Memory Tests

The specific predictions derived for the different views are based on the categorization task. The alternate views could also be evaluated by deriving predictions for memory subsequent to a categorization task. Research in a number of areas in cognitive psychology suggests that memory for information is dependent on the characteristics of the task performed with that information (cf., Craik & Lockhart 1972). Differential recognition memory for pictures and words presented in identical or different forms at test could add to our knowledge about the nature of the categorization process. In fact, a comparison of recognition memory following a specific task (such as categorization) and some baseline (such as a memorization task) provides a means of understanding the unique elements of the specific task. At a broader level, researchers have pointed to differences between conceptual (or semantic) and perceptual processing at study and at test (cf. Weldon et. al. 1989). Hence, recognition memory could be used to study the unique nature of the categorization process.

An issue in recognition is the role of retrieval (Bahrick 1970; Tulving & Thomson 1971). Recognition could be based on familiarity and therefore, retrieval may not play a role (Atkinson & Juola 1973). On the other hand, in certain cases recognition may not be based on familiarity and, therefore, retrieval may play a role (Atkinson & Juola 1973). A study where recognition may have been based on familiarity was conducted by Snodgrass & McClure (1975). In this study, Ss were required to familiarize themselves with a set of pictures and words and instructed to study them with the intent of performing a recognition test. They were then tested on instances where the form of the stimulus at test was manipulated to be either the same or different from its form at study. Their results suggest that, when response bias is taken into account, either, "retrieval of an old concept is equally good when either form is used as a test stimulus or that no retrieval is necessary in recognition memory".

Snodgrass & McClure (1975) used a familiarization technique where Ss went through a study phase and knew beforehand that they were required to complete a recognition test. The present study proposes an unexpected recognition test after a categorization task. The categorization task will entail exposure of instances to Ss for a small amount of time for the purpose of making a semantic decision. Such a procedure would make the argument that recognition is based on familiarity and not on retrieval less plausible. If retrieval is assumed to be necessary for recognition under such conditions, a direct test of the different views of categorization is also possible by examining recognition memory.

Given the assumption that retrieval is required for recognition, the different views would predict differential effects. The "picture" view suggests that input to the categorization process is closer to pictorial than to verbal form. This suggests that a word in the process of categorization, may be transformed to an input which is closer to its pictorial equivalent. In other words, the categorization process may involve the activation of certain elements in memory that are closer to the pictorial form of a concept. Therefore, to the extent that such activation is involved in an initial categorization task, subsequent recognition requiring retrieval may involve these same elements. Retrieval may therefore be facilitated by a pictorial target at recognition due to its hypothesized closeness to elements activated during the initial categorization. Given the assumption that the unexpected recognition task requires retrieval of elements activated during categorization, this view would hypothesize that a *test* stimulus in *pictorial* form would *facilitate* the reactivation of these elements in order to make a recognition decision. Therefore, a facilitation of reaction time is hypothesized for test stimuli in pictorial form (Table 1). Similarly, enhanced accuracy is expected for pictorial test stimuli due to a minimum of transformations required in order to match the test stimuli with information available through the initial process of categorization. Following a similar line of reasoning, the "verbal" view would predict faster and more accurate recognition of verbal stimuli. The "amodal" view would predict equal recognition (in terms of RT and accuracy) of pictures and words.

It should be noted that we do not hypothesize effects for any direct comparison of contrasts of different original modality (i.e., wp vs pw). The line of reasoning presented here does not rule out any effect due to perceptual fluency in recognition (or recognition on the basis of familiarity, Snodgrass (1984)) but argues for the important role of retrieval due to the nature of the categorization task. This argument is one of the degree or extent to which a factor, such as retrieval, plays a role in recognition memory as compared to perceptual fluency. Inasmuch as perceptual fluency is affected by the modality of the original stimulus (due to, say, the picture superiority effect, (Nelson et al. 1976)), comparisons are made keeping the level of this factor a constant (i.e., comparisons of pp and pw *or* wp and ww respectively). Hypotheses are generated for the verbal view using similar arguments.

In summary, we intend to infer theoretical support from the results of the categorization task and the subsequent recognition memory task. In conducting the tests, we employ two criteria for response times (cf., Snodgrass 1984), those of equality or inequality of base times (for the categorization task) and facilitation effects (for the recognition task).

EXPERIMENT 1

Overview

The first experiment consisted of exposing Ss to a stimulus (which is a picture or a word). This was followed by exposure to another stimulus (which is again a picture or a word). Ss were required to indicate whether these pairs of stimuli belong to the same category or not. A sufficient interstimulus interval was provided to facilitate Ss activating a category when exposed to the initial stimulus. The interstimulus interval provided was expected to be sufficient for Ss to have activated a category on exposure to the initial stimulus, with the rationale for the chosen interval provided subsequently. Therefore, the possibility of Ss observing both stimuli and making a judgment without accessing the superordinate (for example, looking at perceptual features and making a judgment) was minimized. If such a process were operating, then latency differences could be attributed to factors other than the relative speed of access to a category and this is undesirable for the present study. Upon exposure to the subsequent stimulus, Ss were required to access the previously activated category and make a Yes/No judgment. Subjects were 27 undergraduates at a Midwestern University who volunteered for participation in the experiment and were paid for their participation.

Stimuli

The stimuli were chosen from the standardized set of pictures and their names developed by Snodgrass & Vanderwart (1980). The instances used were chosen randomly from several categories such as animals, human organs, clothing, fruits, vehicles, sporting items, insects, etc. The mean complexity ratings of the pictures and words were 2.97 and 3.02 out of 5, respectively, according to norms developed by Snodgrass and Vanderwart (1980). The mean image agreement of pictures (a measure of the degree of agreement on the concept represented by an image) was 3.59 out of 5 and the mean familiarity of words and pictures used were 3.49 and 3.51 out of 5, respectively, according to norms developed by Snodgrass and Vanderwart (1980). Moderately high levels of image agreement and familiarity were used in order to facilitate activation of a superordinate category when exposed to the initial stimulus. Familiar instances of a category were used so that Ss were less likely to make an incorrect category decision when exposed to the initial stimulus and, as a consequence, be less likely to recategorize the initial stimulus after exposure to the subsequent stimulus.

Procedure

The experiment was carried out with the use of Macintosh computers. Ss were asked to read instructions before the start of the experiment. These instructions reflected the earlier concerns about the desired process and were repeated verbally to the Ss. The essence of these written and verbal instructions was to direct Ss to identify a category when they were exposed to a stimulus and to compare the subsequent stimulus to the category identified earlier. The Ss were instructed to respond quickly but without compromising on accuracy.

Each (initial) stimulus was presented on the computer screen for 750 ms and this was followed by a blank screen for 500 ms (Figure B, panel a). Words were presented in lower–case bold letters on the center of the screen, while pictures were from the standardized set discussed earlier. The inter–stimulus interval in terms of the time allowed for activation of the category was, therefore, a total of 1250 ms. This particular inter–stimulus interval was chosen for the following reason. Anderson & Reder (1974) present estimates of reaction times for category generation and instance reading. They concluded that the time taken to generate a category is about 400 ms after subtracting out the time taken to encode an instance and the time taken to provide a response. They also showed that an instance reading time is about 750 ms. Based on these findings an inter–stimulus interval of 1250 ms was chosen. The instance was exposed for 750 ms to allow encoding and a blank screen was shown for 500 ms. The blank screen was provided to encourage the Ss to activate a category and also provide an interval between instances in order to prevent a direct perceptual comparison of a pair of instances.

Insert Figure B about here

The inter–stimulus interval was followed by the presentation of the subsequent stimulus. After the Ss had provided a response, a masked screen was shown for 3s to mark the end of a trial. The masked screen also served to end processing of stimuli in the previous trial and to signal the next trial. The Ss were presented with forty eight trials in all with the first eight treated as practice trials for purposes of analyses. There were an equal number of trials representing each possible combination of pictures and words (i.e., ten each of picture–picture(pp), pw, wp, and ww). The correct response was a "Yes" for half the trials and "No" for the other half of the trials.

This part of the experiment was followed by a numerical distractor task that lasted for two minutes. The distractor task was followed by a provision of instructions for the recognition test. Similar instructions on speed–accuracy were provided as in the categorization task. Ss were presented with forty trials and were required to provide a Yes/No response to indicate whether they recognized the stimuli as having been presented in any form (picture or word) in the initial part of the experiment. Twenty of these trials were randomly selected from the set of stimuli presented earlier in the study. These twenty trials consisted of ten trials that were pictorial and ten trials that were verbal when presented in the previous categorization task. These trials were presented at recognition, such that half were in a different modality than at study while half were in the same modality. An additional twenty trials consisted of foils of ten pictures and ten words. On completion of this test, Ss were paid for their participation and debriefed.

Manipulation Test

The experimental procedure outlined above is different from the experimental procedures employed in past research. The argument was made earlier that the procedure used here is more appropriate for studying categorization than experimental procedures used in the past. The critical element to any procedure designed to study categorization is that it involve access to an underlying category. A pretest was performed on the proposed experimental procedure to investigate whether this procedure involves category access during the inter–stimulus interval that is provided between category instances.

The pretest required subjects to perform the proposed experimental task outlined earlier with one difference. Subsequent stimuli for 25% of the trials were replaced by a question regarding the category to which the initial stimulus belonged (such as "Is this an animal?"). These questions (referred to as category inserts) required a similar response (i.e., Yes/No) as the rest of the trials. The correct response to half the inserts was "Yes" and to the other half was "No". Ss were forewarned that such questions "may be posed occasionally" but were asked to focus on the primary task. Therefore, the random inserts were not expected to intrude on the primary task since Ss were instructed to focus on the primary task and the response to these random inserts (i.e., Yes/No) was similar to the rest of the trials. The purpose of such a pretest was to obtain an estimate of the response time to such random inserts in the proposed experimental task. This estimate could be compared to response times for random inserts in a task where categorization is not involved and in another task where category access is required. If the target task involves category access, then it is argued that the response time for such random inserts would be less than the corresponding response time for random inserts in a non–categorization task and equal to the corresponding response time for inserts in a second categorization task.

The pretest, therefore, involved two other tasks; a letter identification task and a typicality rating task (Rosch, 1975). Ss in the letter identification task were exposed to words and required to provide a Yes/No response as to whether or not a particular position in a word was occupied by a specific letter. This task does not involve access to the category corresponding to the word, but only recognition of letters in a word. The typicality rating task involved exposure of Ss to a word or a picture. Ss were instructed to identify the underlying category to which a word or a picture belonged and provide a Yes/No response as to whether the word or picture was a typical member of the identified category. This task was chosen since a category has to be accessed in order for it to be performed. The category inserts were placed at the end of 25% of the trials for both of these tasks. The inserts, as well as their positions within the tasks, were similar to the inserts in the target experimental task. In the letter identification task, a category would have to be accessed after exposure to the category insert. In the typicality rating task, a response to the category insert would require a comparison with a category that had already been accessed in order to perform the rating task. Therefore, the response times to the category inserts provide base–lines to which the corresponding response times from the target experimental task can be compared.

The pretest was carried out with fifteen subjects, five subjects performing each of the tasks. The reaction times for correct Yes and No responses to inserts were analyzed separately. A oneway analysis of variance was performed on the mean response times across the three levels of the task variable and a significant main effect was found for correct Yes responses ($F(2,12) = 17.8$, $p < .001$) and for correct No responses ($F(2,12) = 32.7$, $p < .001$). Pairwise contrast tests were performed for differences between the tasks. Mean correct *Yes* times were; 918ms, 1185ms, and 2519ms, respectively, for the target experimental task, the typicality rating task, and the letter identification task. Mean correct *No* times were 942ms, 1128ms, and 2374ms, respectively, for the target experimental task, the typicality rating task, and the letter identification task. The typicality rating task and the target experimental task were found to have significantly smaller response times than the letter identification task for the correct Yes responses ($t=5.6$, $p<.001$ and $t=4.6$, $p<.001$, respectively) and for the correct No responses ($t=7.4$, $p<.001$ and $t=6.5$, $p<.001$, respectively). As expected, the differences between the typicality rating task and the target experimental task were not significant. Therefore, it is concluded that the experimental context in the target experimental task leads to category access at the end of the inter–stimulus–interval at a level comparable to a typicality task involving categorization, and satisfies the criterion necessary for introducing this methodology into our experimental procedure.

RESULTS

The first step in the data analysis was to exclude reaction times which were beyond two standard deviations from the mean for each subject, as recommended by Srull (1984). A tight criterion of two standard deviations was employed to ensure that the desired process was taking place.[1] Such a criterion was required, since it was necessary to exclude data due to Ss waiting for the subsequent stimulus before making a category decision and, hence, taking more time to do so. The means of Ss for each combination of factor levels (pp, pw, wp, and ww) were calculated for the following dependent variables; correct Yes and No RTs for categorization,[2] correct Yes and No RTs for recognition, hits and false alarms for recognition.[3]

Reaction Times – Categorization Task

An analysis of variance of the data from the correct Yes RTs for categorization and the correct No RTs was performed and the results are presented alongside the hypotheses in Table 1, while the individual RTs are presented in Table 2. A similar analysis was performed on correct Yes RTs for recognition and the number of hits (Table 2). The recognition data was further analyzed for false alarms and a d' value (Murdock 1982) was computed for each of the conditions (Table 2). The d' parameter developed using signal detection theory (Banks 1970) is the distance between the means of the noise distribution (based on the false alarm rate) and the signal distribution (based on the hit rate). This parameter takes into account bias on the part of Ss to respond with a "Yes" more frequently for one mode of presentation over another.

Insert Table 2 about here

---

[1] The outlier analysis for Yes RTs for categorization led to the deletion of 3.3% of the trials. An ANOVA performed on the number of outliers suggested that the number of outliers were not significantly different across conditions (pp, pw, wp, and ww). The outlier analysis for No RTs for categorization led to the deletion of 6.5% of the trials. An ANOVA performed on the number of outliers suggested that two conditions (wp and ww) had a proportionately larger number of outliers. As will be reported subsequently, consequent to deletion of outliers, wp and ww were still found to have significantly higher RTs than the other two conditions.

[2] 4.9% of the trials for the categorization task had incorrect responses.

[3] Due to inadvertently miscounting the number of trials in each condition, in expt. 1, only 4 (instead of 5) trials were used for purposes of analyses for the pw condition for Yes RTs in categorization and the pp condition for Yes RTs in recognition. One ww trial in recognition utilized a target stimulus that was presented earlier in the recognition task. Analyses with and without this trial led to qualitatively similar results. The pw condition for Yes categorization contained 4 (instead of 5) trials. Inclusion of a pw trial from the practice set (i.e., first 8 trials) led to qualitatively similar results.

A 2 (mode of initial stimulus) by 2 (mode of subsequent stimulus) factorial ANOVA was performed on the means from the correct Yes response times for the categorization task. The main effect of the mode of the initial stimulus was significant with faster categorization of pictorial stimuli ($F(1,26) = 4.29$; $p<0.05$), the main effect of the mode of the subsequent stimulus was also significant with faster categorization of pictorial stimuli ($F(1,26) = 15.95$; $p<0.001$), and the interaction was marginally significant ($F(1,26) = 3.28$; $p<0.09$).

The results are presented alongside the hypotheses in Table 1 and the mean RTs for categorization are presented in Table 2. The amodal view had two predictions of which only one was confirmed. The "picture" view was supported for three of the six predictions and directional support was found for two additional predictions. The "verbal" view was supported for only one of the six predictions. It should be noted that the predictions that were confirmed for the amodal and verbal views were consistent with the picture view as well. Hence, the picture view appears to be supported to a relatively greater degree than the other views. While the amodal view is rejected for only one of its predictions it should be noted that it makes only two predictions in all and, therefore, does not account for a majority of the results.

In an effort to uncover potential findings which may have been hidden due to outliers, the data on Yes RT for categorization was further analyzed in the following manner. If Ss had one or more means for a condition which were greater than 2 standard deviations from the average of the means for all subjects, data for such subjects were deleted from the analysis. This procedure resulted in the deletion of data for four subjects. A comparison of deleted data with included data suggested that the number of trials that were considered outliers was higher for the deleted data, but not significantly so (Mean number of trials greater than 2 standard deviations = 5 & 2.2 respectively for deleted and included data; $t = 2.25$; $p < 0.12$). The analysis of included data led to the same three predictions providing support for the picture view. In addition, two other predictions were supported (i.e., wp < ww ($F(1,22) = 3.64$, $p < 0.07$); wp < pw ($F(1,22) = 7.64$, $p < 0.05$). Mean RTs for pp, pw, wp, and ww, were 656, 770, 711, and 757 ms., respectively. In order to be consistent across analyses, similar deletions were performed on all data based on response latencies and the results are indicated along with each analysis.

As explained previously, no assumptions were made about the relationship between the access times into the imagistic and acoustic stores, respectively. It could be argued that the amodal view, with the additional assumption of faster access into the imagistic store than the acoustic store, (i.e., a < a' in Figure A) could make the same predictions as the picture view. Hence, the complete pattern of findings across both categorization and recognition tasks needs to be considered before an explanation can be supported.

Recognition Tests

Similar ANOVAs were performed for the correct Yes RTs and d' values for recognition. For the ANOVAs on Yes RTs, the main effect of the mode of the stimulus at *categorization* was significant with faster recognition of pictorial stimuli ($F(1,26) = 21.35$; $p < 0.001$). The main effect of the mode of the stimulus at *recognition* was also significant with faster recognition of pictorial stimuli ($F(1,26) = 4.24$; $p < 0.05$), and the interaction was significant ($F(1,26) = 15.81$; $p < 0.001$). For the ANOVA on d' values, the main effect of the mode of the stimulus at *categorization* was significant with more accurate recognition of pictorial stimuli ($F(1,26) = 52.02$; $p < 0.001$). In addition, the main effect of the mode of the stimulus at *recognition* was significant with more accurate recognition of pictorial stimuli ($F(1,26) = 10.70$; $p < 0.01$), and the interaction was non–significant. Correlations between response times and hits were negative, thereby ruling out a speed–accuracy trade–off as a potential explanation (the correlations across subjects were –0.19, –0.05, –0.05,and –0.19 for the pp, pw, wp, and ww conditions respectively).

The results of the recognition test are presented alongside the hypotheses in Table 1. The mean RTs and accuracy across different conditions are presented in Table 2. The results show that one prediction based upon the picture view was supported in terms of *recognition speed* and both predictions were supported in terms of *accuracy*. Qualitatively similar results were found following analysis after deletion of data for subjects using the procedure employed for Yes RTs in categorization. No other view finds a comparable level of support for its hypotheses. In light of the evidence from experiment 1, it appears that the categorization process requires input information which is more easily transformable from pictures than from words.

In addition to analyses of Yes responses, an ANOVA was performed on No RTs for categorization. Results for No responses show a significant main effect for the mode of the initial stimulus with faster responses to pictorial stimuli ($F(1,26) = 79.58$; $p < 0.001$), a non–significant main effect for the mode of the subsequent stimulus, and a non–significant interaction. (Qualitatively similar results were obtained for analyses after deletion of data for selected subjects.) The results were quite different from the results for Yes RTs for categorization (to summarize, pp = pw < wp = ww; see Table 2) suggesting that a No response may involve a different process and require additional hypotheses. From the results, it appears that the mode of the initial stimulus is most important for a No response, with faster responses when the initial stimuli are pictorial in form.

EXPERIMENTS 2 AND 3

In order to replicate the findings in the first experiment, we conducted two additional experiments by incorporating variations of the categorization task and the memory test. Since many aspects of these experiments follow the first experiment, only variations in the methods will be noted here.

Procedures

In the second experiment subjects were required to assess the typicality of stimuli, a task employed in the manipulation test that also involved category access. The stimuli used were chosen from the set developed by Snodgrass & Vanderwart (1980). Again, the instances were from a range of categories similar to the first experiment. According to norms developed by Snodgrass and Vanderwart (1980) the mean complexity of words and pictures used in this task was 2.77 and 2.92 out of 5, respectively. The mean familiarity of words and pictures was 3.48 and 3.61 out of 5, respectively, and the mean image agreement of pictures was 3.78 out of 5. Subjects were exposed to a stimulus in pictorial or verbal form and asked to rate the extent to which an instance was a "good" example of the category to which it belonged (Figure B, panel b). Below each stimulus on the computer screen was a seven point rating scale which was similar to the scale used by Rosch (1975). The instructions reflected the content of the procedure employed by Rosch (1975) except that Ss were not provided with a category name. This was to prevent any bias towards verbal information as discussed by Snodgrass (1980). Subjects were required to identify a category and this task was expected to be fairly accurate due to the use of familiar instances as stimuli. Subjects were exposed to forty–eight trials, with the first eight being excluded from subsequent analysis. These trials consisted of twenty–four pictures and twenty–four words in mixed order. The reaction time for each response was the time from beginning of exposure of a stimulus until a key was hit to rate the typicality of an instance. This part of the experiment was followed by a two minute numerical distractor task.

The distractor task was followed by a recognition test which was similar to the recognition test in experiment 1.

The typicality task was chosen in order to introduce a variation in the basic categorization task and replicate our previous findings. This task required Ss to categorize an instance and then rate it on its degree of membership in the identified category. Hence, it incorporated the basic element of category access, but at the same time is different from the tasks employed earlier (such as a same–different task or a Yes–No judgment) in that it requires a judgment of typicality. This task was used with favorable results in the manipulation test reported earlier. The predictions for experiment 2 are the same as the predictions for experiment 1 and are presented in Table 1.

Experiment 3 was performed in order to evaluate the competing explanations using a different dependent variable. A key assumption discussed earlier was that the recognition task required retrieval given the nature of the primary task and the unexpectedness of the recognition test. In order to provide a further test for retrieval a recall task was substituted for the recognition task in experiment 3 (Figure B, panel c). Experiment 3 was similar to experiment 2 in all other respects. Subjects were required to list out the items they could recall from the typicality task and also to report the form (i.e pictorial or verbal) in which an item was accessed from memory (i.e., "came to mind"). The predictions for Experiment 3 are the same as the predictions for Experiment 1 and are presented in Table 1.

RESULTS

Experiment 2–Typicality/Recognition

The results of experiment 2 are presented in Table 1 and the mean RTs and accuracy for different conditions are presented in Table 2. The RT for typicality ratings of pictures and words were not significantly different (Mean for pictures =3497 ms and Mean for words =3374 ms). The large RTs for the typicality rating task as compared to the categorization task used in experiment 1 were indicative of the time taken to rate instances on their judged typicality. The order of RTs for the rating task allowed the Ss to be exposed to the stimuli for longer periods of time than in experiment 1. ANOVAs were performed for the correct Yes RTs and d' values for recognition. For the ANOVAs on Yes RTs, the main effect of the mode of the stimulus at *categorization* was significant, with faster recognition of pictorial stimuli ($F(1,24) = 24.05$; p <

0.001). The main effect of the mode of the stimulus at *recognition* was also significant with faster recognition of pictorial stimuli ($F(1,24) = 5.48$; $p < 0.05$), and the interaction was significant ($F(1,24) = 15.43$; $p < 0.001$). For the ANOVA on d' values, the main effect of the mode of the stimulus at *categorization* was significant with more accurate recognition of pictorial stimuli ($F(1,28) = 10.28$; $p < 0.01$). However, the main effect of the mode of the stimulus at *recognition* was non–significant, and the interaction was non–significant. Correlations between response times and hits were negative, thereby ruling out a speed–accuracy trade–off as a potential explanation (the correlations across subjects were –0.09, –0.03, –0.16, and –0.09 for the pp, pw, wp, and ww conditions respectively).

In terms of the accuracy of recognition, the picture view was supported for one of its predictions. In terms of the reaction time results for recognition, this view was also supported for one of its predictions. Again, no other view received better support than the picture view. Qualitatively similar results were obtained after deletion of data on selected subjects, as in experiment 1.

Experiment 3–Typicality/Recall

A 2 (mode of stimulus at exposure) by 2 (mode of stimulus reported at recall) factorial ANOVA was performed on the number of items accurately recalled. A significant main effect for mode at *exposure* was found with more accurate recall of pictorial stimuli ($F(1,32 = 59.38$; $p < 0.001$). A significant main effect was found for mode of stimulus reported at *test* with more accurate recall of stimuli reported to be in pictorial form ($F(1,32) = 87.71$; $p < 0.001$). A significant interaction was also found ($F(1,32) = 210.83$; $p < 0.001$).

The results of experiment 3 are presented in Table 1 and the results of the analysis of recall data are presented in Table 2. In interpreting these results, it should be mentioned that the recall measure which required a self–report of modality of retrieval is open to the usual criticisms of self–report measures (Nisbett & Wilson 1977). Only one of the two predictions of the picture view was supported. As is evident from Table 2, there is a predominant main effect for retrieval in pictorial form and better recall of pictorial stimuli (the latter finding has been obtained in past research (Erdelyi and Becker 1974)). One possible explanation for superior retrieval in pictorial form would be a general tendency of Ss to guess in this direction, thereby inflating the relevant cells. However, an analysis of false recalls (i.e., recall of items which were not presented to the Ss) showed no bias in either direction.

DISCUSSION OF RESULTS

In summarizing the results, it is useful to look at the findings across experiments. With respect to Yes RTs of categorization, there appears to be support for the picture view except for the low RTs for pp compared to wp. For No judgments, however, the key factor appears to be the mode of the initial stimulus rather than the subsequent stimulus. In addition, recognition memory following categorization can be compared to recognition following a memorization task. For the latter, Snodgrass and McClure (1975) found a picture superiority effect in terms of accuracy (pp = pw > wp = ww). Our results obtained for recognition following categorization selectively deviate from these prior findings. First, is the finding across experiments and dependent variables of pp < pw (in terms of RT) and pp > pw (in terms of accuracy). Hence, following categorization, as distinct from memorization, there appears to be an effect of mode at recognition, especially when the stimulus is initially pictorial in form. In comparing pw and wp, pw is recognized marginally faster than wp ($F(1,26) = 2.94$; $p < .10$) and is equal to wp in terms of accuracy (experiment 1). In experiment 2, pw is recognized faster than wp ($F(1,24) = 4.93$; $p < .05$) and as accurately as wp. Hence, the superiority of pictures in terms of accuracy was not found (Snodgrass and McClure, 1975); report pw > wp for accuracy). Arguably, facilitation for pictorial targets following categorization may have nullified the picture superiority effect.

In comparing wp and ww, the only result hypothesized according to the picture view which received support is that wp > ww (experiment 1). It appears that, while the categorization task may have an effect on subsequent recognition memory, this effect is moderated by the mode of the stimulus being categorized. Clearly, the results relating to wp and ww cannot be explained by the picture view in its simple form. Finally, in comparing pw and ww, pw = ww in terms of RT in both experiments. In contrast, when examining accuracy, pw > ww (experiment 1) versus pw = ww (experiment 2). Again, the picture superiority effect was not consistently found following a categorization task. This is in line with the notion that, following categorization, the mode of the target stimulus at recognition may be important.

In terms of recall, the results could be summarized as follows (pp > ww > wp > pw). Apart from the higher recall of pictorial information, there was a strong effect of recall in the same mode as at exposure (i.e., pp and ww were high). Hence, ww was found to be greater than wp. It is possible that the requirement that subjects report the mode in which they recalled information may have led to a greater degree of reconstruction of the original stimulus which in turn may have lead to identification of the original mode of information. Hence, the more appropriate comparisons may be between wp and pw as they interact with recall following categorization versus memorization tasks. It should be noted that, even though a picture superiority effect was found, wp was greater than pw (with false recalls not being significantly greater for one mode over the other).

GENERAL DISCUSSION

A Process Model of Picture-Word Categorization

The overall pattern of results for the three experiments point toward a categorization process with information requirements which are more compatible with pictorial than with verbal stimuli. In terms of a semantic network with spreading activation (Anderson & Bower 1973), it is possible that pictures activate their visually salient featural nodes either *prior* to or *simultaneous* with the activation of their concept nodes. Words, however, activate their concept nodes and this activation may *subsequently* spread to featural nodes (since words bear no resemblance to their features they may not be able to access feature nodes directly). In other words a picture of a concept is also a picture of its perceptual features.

While traditional accounts of a picture advantage attribute time differences to encoding operations *prior* to semantic access, such as, greater picture distinctiveness and acoustic–phonemic coding of words (Friedman and Bourne 1976; Pellegrino et al. 1977), our model suggests a picture advantage in semantic *access*. In this respect, common coding views of semantic memory have been criticized by dual coding theorists (te Linde 1982) for predicting a constant picture advantage in semantic decisions. Paivio & te Linde (1980), however, have shown that a picture advantage is eliminated for semantic decisions along certain dimensions. Our model, while predicting a picture advantage in semantic decisions along visually perceivable dimensions, does not suggest a *constant* picture advantage for *all* semantic decisions. We suggest that the occurrence of a picture advantage is dependent on the task at hand and the extent to which **visually perceivable** information can be utilized for performance of the task. Hence, a picture advantage in semantic tasks could occur due to the elements of semantic memory that can be activated *directly* by the visually salient information available in pictures.

A number of alternate explanations of the picture–word categorization process were evaluated in this paper. Based on the issues discussed earlier, a process model of picture–word categorization is presented (Figure C). The thrust of this model is that pictures **simultaneously** access both the concept node and visually salient feature nodes. This stage of processing leads to a picture advantage in categorization. Therefore, pictures are preprocessed in some sense for purposes of categorization. In comparison, a word initially accesses its concept node which **subsequently** activates feature nodes (which are not necessarily visually salient). *Simultaneous* versus *sequential* activation of featural nodes explains differential transformation of picture versus word stimuli at exposure. Subsequent recognition memory is affected by the nature of the categorization task and whether it involves the activation of visually salient features.[4]

Insert Figure C about here

Based upon the model presented, we can explain our findings with respect to the categorization task. The reaction time finding that pp < wp would be explained by advantages accruing to both pictures in the pp combination activating an overlapping set of features, thereby hastening the categorization process. Hence, the strong categorization finding that pp is faster

---

[4] We speculate that differences in latencies between pictures and words may also be due to differential processes *in* semantic memory rather than just differential access *into* semantic memory. The comparison process to categorize an instance could be either holistic (non analytical) or featural (analytical) in nature (Smith and Nelson 1984). This depends on the extent to which stimuli are processed by separate stimulus dimensions versus overall similarity among stimuli. Holistic processing is expected to be speedier than analytic processing (Smith and Nelson 1984).

A typical finding in developmental psychology is that very young children tend to categorize in a holistic fashion and tend to use perceptual attributes as the basis for categorization (cf., Shepp 1978). It is possible that these two phenomenon are theoretically related to each other. Pictures, by presenting visually perceivable attributes in simultaneous fashion, may invoke holistic processing of information. In other words, pictures present a combination of features in simultaneous fashion and are, therefore, likely to be processed in a holistic fashion. Smith and Nelson (1984) show that adults tend to use overall similarity in categorization under conditions such as time pressure and task load. They argue that separable dimensions (such as size and brightness) have both analytic and holistic modes of classification available but integral dimensions (such as saturation and brightness) have only a holistic mode available. The concepts of integrality and separability are recognized by the authors as being a matter of degree. We suggest that pictures may facilitate the perception of separable features as being integral. In doing so, pictures may facilitate a holistic mode of classification. Adults appear to use a similarity mode of classification under certain conditions such as "stimulus features (complexity, integrality),...." (Smith and Nelson 1984). We propose that the modality of information (by increasing or decreasing perceived integrality) may be one such factor which may lead to the use of different classification modes. Therefore, in the presence of a stimulus with separable dimensions in pictorial form, people may have a tendency to use a holistic mode of processing.

than other combinations. Subsequent effects on recognition can be explained in the following manner; pp is recognized faster and more accurately than other combinations due to the *overlapping set of features* that may be activated. The recognition of wp may be facilitated to the extent that the categorization of words leads to the activation of feature nodes, some of which may be *visually* salient. Hence, facilitation may be dependent on the amount and type of activation of feature nodes.

As mentioned earlier, the No response in categorization appears to require a different process than the Yes response. In particular, the mode of the initial stimulus appears to play an important role in determining response speed. It is possible that initial stimuli in pictorial form, through simultaneous activation of featural nodes, facilitate a No response. This could be due to an easier determination of *lack of overlap* between featural nodes of stimuli from *different* categories. These post–hoc explanations, although supported by our findings, are tentative, at best, and additional theory development, testing, and refinement is required.

Within the model and this research, we make an implicit assumption that a comparison process is necessary for categorization to take place. One issue that has not been addressed is the possibility that categorization does not necessarily involve some comparison process, but rather is oriented towards retrieval (Holyoak and Glass 1975). Collins and Loftus (1975) argue that both of these processes may be used in categorization. Lorch (1981) shows support for the hypothesis that people can vary their dependence upon these approaches. The results of experiments conducted by Lorch (1981) suggest that a retrieval strategy is used when the items in the task are meaningfully related whereas a comparison strategy is used when meaningfully unrelated items occur in the task. This study used both unrelated and related instances in the categorization task in experiment 1, thereby suggesting that subjects may have used a comparison process. Lorch makes the distinction between semantic overlap and relation strength; the former being "the extent of shared associations between the subject and predicate concepts" and the latter being "the strength of the most accessible subject–predicate connection which is sufficient to determine a response" (Lorch 1981). In a process involving the retrieval of category information, a picture advantage in categorization by retrieval could accrue from multiple paths to the category node due to the simultaneous activation of concept and featural nodes. In the case of word categorization by retrieval, it should be noted that the activation of featural nodes by the concept node is only a by–product of processing and is not a necessary step in categorization (as opposed to a comparison process where this step is essential). A word would be categorized by a retrieval process based on the "relation strength" between the concept and its category. Thus, any model of picture–word categorization would necessarily have to account for both *retrieval* and *comparison* processes, taking place to different degrees, and which include effects due to *relation strength* and *semantic overlap.*

In using a comparison strategy, subjects appear to base their judgments in sentence verification tasks on the degree of "semantic overlap" between the subject and the predicate (Smith et al. 1974). Lorch (1981) suggests that the degree of semantic overlap may be judged on the basis of the number of connections between the subject and the predicate and/or the sum of the levels to which connections between the subject and the predicate are activated. This argument is similar to the explanation provided earlier for the finding that picture–picture combinations require lesser RTs in the categorization task than word–picture combinations. Therefore, the notions of relation strength and semantic overlap appear to offer useful explanations for findings in the categorization of pictures and words.

At this point in the discussion it is necessary to compare the explanations of the model presented above with several alternate views. Snodgrass (1980, 1984) suggests that the locus of the picture advantage may be at Level III processing due to visual appearance rather than meaning. This explanation is different from the view that a picture advantage in categorization occurs *prior* to semantic access in that it suggests picture categorization *without* semantic access. Such an approach may be applicable to category decisions for picture–picture combinations, but not for other combinations of stimuli. It should be noted that we consistently obtained the lowest RTs for picture–picture combinations. As discussed earlier, the explanation derived from our process model is based upon the degree of *semantic overlap* generated by a pair of pictures *within* the same category. On the other hand, visual similarity *across* categories interferes with this process by accessing an overlapping set of instance to instance feature nodes or by preventing access to unique feature nodes. In contrast, visual dissimilarity *across* categories leads to access of sets of feature nodes with little or no overlap between instances (leading to small No RTs for categorization). The model presented in this paper postulates an effect of pictures due to visually salient features activating their respective stored featural nodes, and visual similarity across categories interferes with this visual salience. Although, the processing of features is common to both explanations, our model places more emphasis on differential category access for pictures versus words and the importance of feature overlap for exemplar to exemplar comparisons

IMPLICATIONS AND FUTURE DIRECTIONS

This research suggests several implications for the study of consumer information processing. Our findings and model provide empirical support and a theoretical rationale for speedier and less effortful categorization of pictures versus words. Implications for these findings will be discussed in terms of their interface with the categorization of products, effects on consumer decision making and judgments, and on the assessment of individual differences in the processing of visual versus verbal information.

Cohen and Basu (1987) discuss the practice of "positioning" products to appeal to particular market segments in terms of consumer judgment of a product's category membership. "The outcome of this *process* is not only a particular identification of a product, but *the increased salience of information relevant to that category....* " (Cohen and Basu, 1987, p. 455, emphasis added). These authors present a contingency-based mixed model in which the process varies based upon the (1) nature of category knowledge (rule defined, prototype defined, exemplar defined), (2) nature of the process (analytic versus nonanalytic), and (3) invocation of automatic versus deliberative mechanisms. These factors play a major role in determining an individual's categorization response. Our results suggest that another factor, the structure of the stimulus, also plays a role in determining the type of categorization process that will be invoked and the nature of information that will receive increased salience. With differential category access, picture-based stimuli, such as used in a print ad, over word-based stimuli will be more likely to invoke resultant categorization processes that are more exemplar-based, nonanalytical, and perhaps more automatic. Whereas, more word or verbally based stimuli should be more likely to invoke categorization processes that are more rule-based and analytical (feature by feature) which is representative of more deliberate processing mechanisms. Cohen and Basu (1987) argue that in a given categorization instance there may be a "first cut" using the most efficient means of making a quick judgment. If preattentive automatic processes can create a unitized representation (more "picture-like") overall similarity-based processes become more likely. Smith and Kemler Nelson (1984) also argue that holistic similarity may be the most basic categorization process in that access to components or individual features of a successfully integrated representation may deteriorate at a faster rate. The advantages of less effortful more automatic simultaneous processing of the concept and its features are likely to be even more pronounced as the speed or complexity of the classification task is increased. Thus, this initial access to the category through, for example a picture dominated advertisement, should serve as an important determinant of the types of processes that are invoked during categorization which ultimately have implications for any subsequent decision or judgment.

When making judgments of a product's representativeness in a category (e.g., is a Lexus ES 300 a sports car or a luxury car?) the structure of the advertising stimulus may also affect the type of decision or judgment process that occurs. Loken and Ward (1990) found that frequency of instantiation was a significant predictor of category *typicality*, or the degree to which an instance is perceived to represent a category. In turn, frequency of exemplar retrieval is related to accessibility as determined by such factors as recency, salience, and distinctiveness (Cohen and Basu 1987). Pictures are one factor that have been found to consistently affect these retrieval related determinants (Childers and Houston 1984). Thus, picture-based marketing stimuli, through their faster access to a category, should serve to affect the degree to which consumers judge whether an existing or new product entrant is perceived as a prototypical member of a category. This is particularly likely since pictures also increase instance to instance access within a category.

With faster access to a category the possibility exists that picture-based marketing stimuli may also play a significant role in framing purchase decisions (Tversky and Kahneman 1981). One view of decision making, prospect theory, posits that the choice process has two distinct stages, an editing stage and an evaluation stage (Puto 1987; Maheswaran and Meyers-Levy 1990). Under the editing stage the consumer structures or "frames" the problem into a more simplified representation. This representation serves as a "reference point" that sets up expectations and objectives for the evaluation or choice stage. Information conveyed in a pictorial format should thus be an important determinant of the types of information activated to frame the problem. More accessible attributes that receive increased salience through their pictorial structure would be more likely to provide a foundation for the expectations that lead to the construction of a category prototype or reference point. Related to this is the effect of priming on judgments. Yi (1990) found that prior exposure to a competitive ad can prime certain attributes and subsequently increase the likelihood that consumers interpret new product information in terms of these attributes thereby affecting the evaluation of a product contained in a subsequent ad. These findings would likely be moderated by the dominate modality of the priming ad. Picture-dominated ads would be more likely to create s stronger priming affect and in turn have perhaps a more enduring effect than their verbal counterparts.

The modality of the stimulus information may affect not only the type of information that is first accessed or frames a decision, but also the interrelationships between cues contained in the marketing communication, package, or product itself. Holbrook and Moore (1981) examined the extent to which verbally versus pictorial-based stimuli affected the interrelationship among features of sweaters, or the *configurality* of the feature cues. After controlling for mental imaging, Holbrook and Moore found that pictorial presentations increased the number of interactions among features used in judging the sweaters. Thus, pictures were more likely to invoke a more "gestalt" or nonanalytical type of processing. These results are consistent with our model, whereby pictures are more likely to access the sweater category and lead to a *simultaneous* processing of the concept and its descriptive and evaluative features. Our findings thus serve to link the categorization literature with the decision making literature while making an important distinction between the modality of the stimulus upon which the decision or judgment is made.

Results from this study also have implications for the study of individual differences in visual versus verbal information processing. In the past five years several different scales have been developed in the marketing literature for the study of individual differences. Among these are scales to measure individual preferences for visual versus verbal information (Holbrook et. al. 1984; Childers, Houston, and Heckler 1985). The latter uses a verbal descriptor whereby the respondent indicates whether they agree or disagree with scale item. In contrast, Holbrook et. al. use an approach whereby a word or picture is first presented and then a related concept is provided in both a verbal and pictorial form. The verbal and pictorial descriptors serve as anchors along a 7 point semantic differential type scale. For example, the first concept might be "work" and the scale descriptors would be the word "factory" and a picture of a factory. The respondents task is to indicate whether, and to what degree, the response to "work" that "comes to mind" is more "verbal-like" or more "picture-like". Repeating this procedure across ten items subsequently leads to the classification of an individual as a "visualizer" or a "verbalizer". Comparing this procedure to Figure B (panel a) shows the similarity between the Holbrook et. al. scale and the underlying process of first accessing the category (e.g., work) and then making a judgment about the category instance to instance association. Our results would predict that the speed with which this occurs is dependent upon the modality of the category descriptor and the instances being used as anchors. From experiment 1, reaction times for picture-picture instances were on the average 126 ms. faster than when the instances were presented as word-word. Thus, the first instance that "comes to mind" might interact with the modality of the priming stimulus and a chronic picture-advantage might overwhelm or bias the modality of the scale item anchor that is chosen. If the anchor selected is biased toward this chronic picture advantage then the selection might be more reflective of a task-based factor (modality of the stimulus and response) than of an individual difference factor. Thus, results from our research suggest that further assessment of measures of modality processing preferences should be undertaken to assess their construct validity. Included in this research might be an investigation of an underlying explanation for these processing preferences as it might provide insight into the validity of these self-report scales as well.

The present study has investigated the nature of the categorization of pictures and words. Findings from three experiments support the conclusion that pictures have faster access than words to their categories and promote faster category instance to instance associations. Our model of this process asserts that pictures achieve this advantage through their faster access of the category concept and due to the simultaneous activation of overlapping features between category instances. In closing, it appears that the intersection of the area of research now being pursued on categorization with that of the modality of information presentation presents a fruitful avenue for knowledge development on the interplay between consumer memory, judgment, and decision making.

REFERENCES

Anderson, J.R., and G.H. Bower (1973), Human Associative Memory, Washington D.C.:Hemisphere Publishing Corporation.

Anderson, J.R., and Lynne M. Reder (1974), "Negative Judgments In and About Semantic Memory," Journal of Verbal Learning and Verbal Behavior, 13, 664–681.

Atkinson, R.C., and J.F. Juola (1973), "Search and Decision Processes in Recognition Memory," in Contemporary Developments in Mathematical Psychology,(eds.) Krantz et.al., New York: Academic Press.

Bahrick, H.P. (1970), "A Two Phase Model for Prompted Recall," Psychological Review, 77, 215–222.

Banks, W.P. (1970), "Signal Detection Theory and Human Memory," Psychological Bulletin, 74, 81–99.

Childers, Terry, and Michael J. Houston (1984), "Conditions for a Picture–Superiority Effect on Consumer Memory," Journal of Consumer Research, 11 (Sept.), 643–54.

_____, Michael J. Houston, and Susan Heckler (1985), "Measurement of Individual Differences in Visual vs.Verbal Processing," Journal of Consumer Research, 12 (Sept.),125–34.

Cohen, Joel B. (1981), "The Role of Affect in Categorization: Toward a Reconsideration of the Concept of Attitude," Association of Consumer Research, VIX, A. Mitchell (ed.), 94–100.

_____ and Kunal Basu (1987), "Alternative Models of Categorization:Toward a Contingent Processing Framework," Journal of Consumer Research, 13 (March), 455–72.

Collins, A.M., and E.F. Loftus (1975), "A Spreading–Activation Theory of Semantic Processing," Psychological Review, 82,407–28.

Craik, F.I.M., and R.S. Lockhart (1972), "Levels of Processing: A Framework for Memory Research," Journal of Verbal Learning and Verbal Behavior, 11, 671–84.

Erdelyi, M.H., and J. Becker (1974), "Hyperamnesia for Pictures: Incremental Memory for Pictures but not Words in Multiple Recall Trials," Cognitive Psychology, 6, 159–71.

Friedman, A., and L.E. Bourne (1976), "Encoding the Levels of Information in Pictures and Words," Journal of Experimental Psychology: General, 105, 169–90.

Holbrook, Morris B., Robert W. Chestnut, Terence A. Oliva, and Eric A. Greenleaf (1982), "Play as a Consumption Experience:The Roles of Emotions, Performance, and

Personality in the Enjoyment of Games," Journal of Consumer Research, 11 (Sept.), 728–39.

____, and William L. Moore (1981), "Feature Interactions in Consumer Judgments of Verbal versus Pictorial Presentations," Journal of Consumer Research, 8(1), 103-113.

Holyoak, K.J., and A.L. Glass (1975), "The Role of Contradictions and Counterexamples in the Rejection of False Sentences," Journal of Verbal Learning and Verbal Behavior, 14, 215–39.

Loken, Barbara and James Ward (1990), "Alternative Approaches to Understanding the Determinants of Typicality," Journal of Consumer Research 17(2), 111-126.

Lorch, R.F. (1981), "Effects of Relation Strength and Semantic Overlap on Retrieval and Comparison Processes during Sentence Verification," Journal of Verbal Learning and Verbal Behavior, 20,593–610.

Maheswaran, Durairaj, and Joan Meyers-Levy (1990), "The Influence of Message Framing and Issue Involvement," Journal of Marketing Research, 27(3), 361-367.

Murdock, Bennett B. (1982), "Recognition Memory," in Handbook of Research Methods in Human Memory and Cognition, (ed.) C. Richard Puff, Academic Press.

Nelson, D.L., V.S. Reed, and J.R. Walling (1976), "Picture Superiority Effect," Journal of Experimental Psychology: Human Learning and Memory, 2, 523–28.

Nisbett, Richard E., and T.D. Wilson (1977), "Telling More Than We Can Know: Verbal Reports on Mental Processes," Psychological Review, 84, No.3, 231–79.

Paivio, A. (1971), Imagery and Verbal Processes, New York: Holt, Rinehart, and Winston.

Paivio, A., and D.J. te Linde (1980), "Symbolic Comparisons of Objects on Color Attributes," Journal of Experimental Psychology: Human Perception and Performance, 6, 652–61.

Pellegrino, J.W., R.R. Rosinski, H. L. Chiesi, and A. Siegal (1977), "Picture Word Differences in Decision Latency: An Analysis of Single and Dual Memory Models," Memory and Cognition, 5, 383–396.

Pylyshyn, Zenon (1973), "What the Mind's Eye Tells the Mind's Brain: A Critique of Mental Imagery," Psychological Bulletin, 80, 1-24.

____________, (1981), "The Imagery Debate: Analog Media versus Tacit Knowledge," in Imagery, Ned Block, ed., Cambridge, MA: The MIT Press.

_____ and B. A. Faulconer (1975), "Time to Understand Pictures and Words," Nature (London), 253,437–438.

Puto, Christopher P. (1987), "The Framing of Buying Decisions," Journal of Consumer Research, 14(3), 301-315.

Rosch, E. (1975), "Cognitive Representations of Semantic Categories," Journal of Experimental Psychology:General, 104, 192–233.

Shepp, B.E. (1978), "From Perceived Similarity to Dimensional Structure: A New Hypothesis about Perceptual Development," in Cognition and Categorization, (eds.) E.H. Rosen and B.B. Lloyd, Hillsdale,New Jersey:Erlbaum.

Smith, E.E., E.J. Shoben, and L.J. Rips (1974), "Structure and Process in Semantic Memory: A Feature Model of Semantic Decisions," Psychological Review, 81, 214–41.

Smith, J.D., and D.G. Kemler Nelson (1984), "Overall Similarity in Adults' Classification: The Child in All of Us," Journal of Experimental Psychology: General, 113, No.1, 137–59.

Smith, M.C., and L.E. Magee (1980), "Tracing the Time Course of Picture Processing," Journal of Experimental Psychology: General, 109, 373–392.

Snodgrass, J.G. (1980), "Toward A Model for Picture Processing," in Processing of Visible Language,(Vol.2),(eds.) P.A. Kolers, M.E. Wrolstad, and H. Bouma, New York: Plenum.

_____ (1984), "Concepts and Their Surface Representations," Journal of Verbal Learning and Verbal Behavior, 23, 3–22.

_____ and Phyllis McClure Burns (1975), "The Effect of Repeated Tests on Recognition for Pictures and Words," Bulletin of the Psychonomic Society, Vol.11 (4), 263–266.

_____ and Brian McCullough (1986), "The Role of Visual Similarity in Picture Categorization," Journal of Experimental Psychology: Learning, Memory, and Cognition, 12 (1), 147–154.

_____ and M. Vanderwart (1980), "A Standardized Set of 260 Pictures:Norms for Name Agreement,Image Agreement, Familiarity, and Visual Complexity," Journal of Experimental Psychology:Human Learning and Memory, 6, 174–215.

Srull, Thomas K. (1984), "Methodological Techniques for the Study of Person Memory and Social Cognition," in Handbook of Social Cognition, (eds.) Robert S. Wyer Jr. and Thomas K. Srull, Hillside NJ: L.Erlbaum Associates.

Sujan, Mita (1985), "Consumer Knowledge: Effects on Evaluation Strategies Mediating Consumer Judgments," Journal of Consumer Research, 12 (June), 31–46.

te Linde, J. (1982),"Picture–Word Differences in Decision Latency: A Test of Common–Coding Assumptions," Journal of Experimental Psychology: Learning, Memory, and Cognition, 8, 584–598.

Tulving, E. and D.M. Thomson (1971), "Retrieval Processes in Recognition Memory: Effects of Associative Context," Journal of Experimental Psychology, 87, 116–124.

Tversky, Amos and Daniel Kahneman (1981), "The Framing of Decisions and the Psychology of Choice," Science, 211 (January, 453-458.

Weldon, Mary Susan, Henry L. Roediger III, and Bradford H. Challis (1989), "The properties of retrieval cues constrain the picture superiority effect," Memory and Cognition, 17(1), 95–105.

Yi, Youjae (1990), "The Effects of Contextual Priming in Print Advertisements," Journal of Consumer Research, 17(2), 215-222.

TABLE 1
HYPOTHESES AND RESULTS

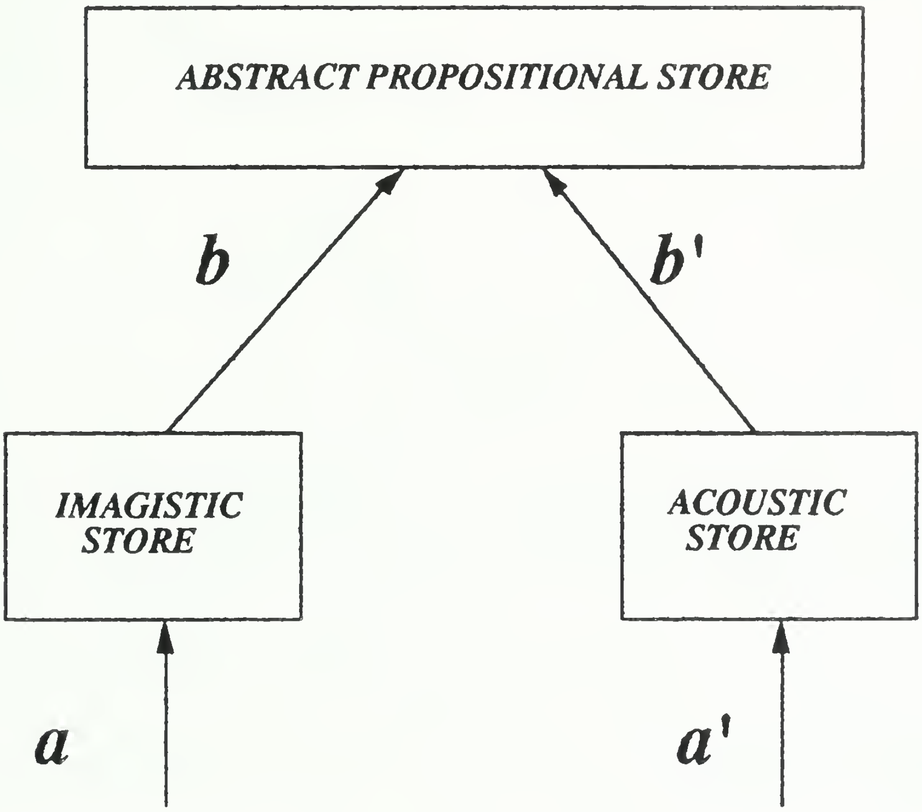
Amodal	Picture	Verbal	Results	F
		YES RT FOR CATEGORIZATION (EXPT. 1)		
PP=WP	PP=WP	PP=WP	PP<WP	7.92*
WW=PW	WW=PW	WW=PW	WW=PW	
–	PP<WW	PP<WW	WW<PP	20.83*
–	WP<WW	WW<WP	WP<WW	1.95***
–	PP<PW	PW<PP	PP<PW	28.11*
–	WP<PW	PW<WP	WP<PW	1.89***
		NO RT FOR CATEGORIZATION (EXPT. 1)		
PP=WP	PP=WP	PP=WP	PP<WP	40.8*
WW=PW	WW=PW	WW=PW	PW<WW	26.58*
–	PP<WW	WW<PP	PP<WW	16.07*
–	WP<WW	WW<WP	WW=WP	
–	PP<PW	PW<PP	PP=PW	
–	WP<PW	PW<WP	PW<WP	46.30*
		YES RT FOR RECOGNITION (EXPT. 1)		
WP=WW	WP<WW	WW<WP	WP=WW	
PP=PW	PP<PW	PW<PP	PP<PW	31.9*
		RECOGNITION ACCURACY (D') (EXPT. 1)		
WP=WW	WP>WW	WW>WP	WP>WW	5.38**
PP=PW	PP>PW	PW>PP	PP>PW	7.31**
		YES RT FOR RECOGNITION (EXPT 2)		
WP=WW	WP<WW	WW<WP	WP=WW	
PP=PW	PP<PW	PW<PP	PP<PW	34.43*
		RECOGNITION ACCURACY (D') (EXPT 2)		
WP=WW	WP>WW	WW>WP	WP=WW	8.13*
PP=PW	PP>PW	PW>PP	PP>PW	3.44***
		NUMBER OF ITEMS RECALLED (EXPT. 3)		
WP=WW	WP>WW	WW>WP	WW>WP	27.41*
PP=PW	PP>PW	PW>PP	PP>PW	255.61*

*$p < .01$, **$p < .05$, ***$p < .20$

TABLE 2
MEANS OF DEPENDENT VARIABLES

	P–P	P–W	W–P	W–W	P	W
Yes RT for Categorization in ms.(Expt. 1)	693	819	772	819	–	–
No RT for Categorization in ms.(Expt. 1)	819	814	958	904	–	–
Yes RT for Recognition in ms.(Expt. 1)	771	989	1072	1011	–	–
Yes RT for Recognition in ms.(Expt.2)	830	1135	1272	1235	–	–
Proportion of Hits in Recog. (Expt. 1)	0.96	0.90	0.79	0.72	–	–
Proportion of False Alarms in Recog.(Expt. 1)	–	–	–	–	0.19	0.22
D' for Recog. (Expt. 1)	4.31	3.24	2.72	1.81	–	–
Proportion of Hits in Recog. (Expt. 2)	0.98	0.85	0.81	0.80	–	–
Proportion of False Alarms in Recog.(Expt.2)	–	–	–	–	0.12	0.09
D' for Recog. (Expt. 2)	5.12	4.21	3.71	3.77	–	–
Ave.Number of Items Recalled (Expt. 3)	10.06	0.42	1.76	4.48	–	–
Average Number of False Recalls (Expt. 3).	–	–	–	–	0.42	0.39

FIGURE A
ALTERNATE VIEWS OF THE CATEGORIZATION PROCESS



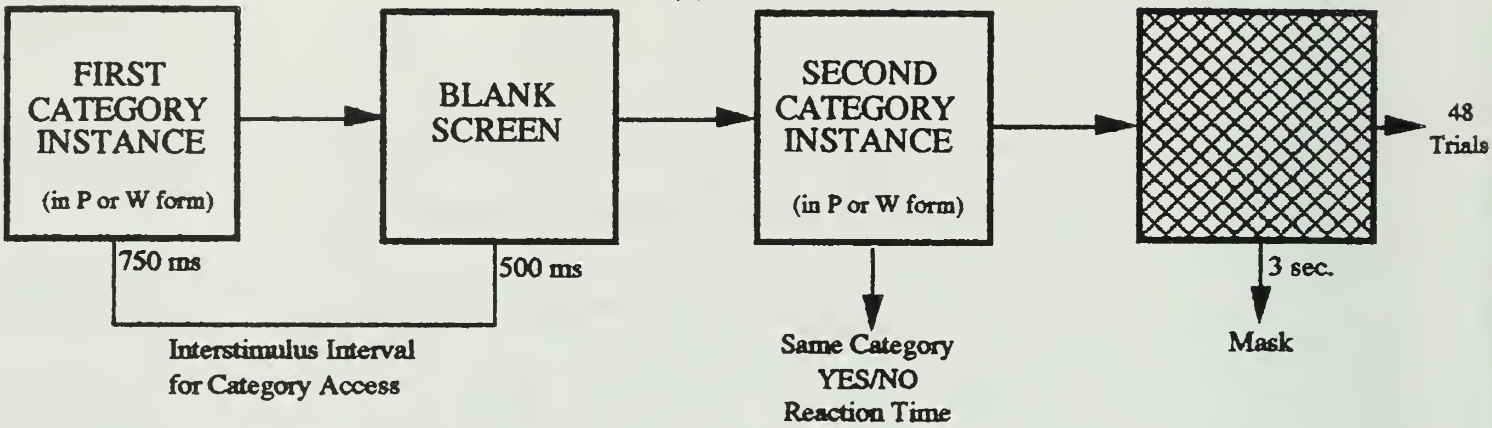
a = access time into imagistic store
a' = access time into acoustic store
b = time taken to categorize input from imagistic store
b' = time taken to categorize input from acoustic store
t = time difference between b and b' due to additional transformations
r = time to produce response after categorization

AMODAL VIEW: b = b'
PICTURE VIEW: b' = b + t
VERBAL VIEW: b' = b + t

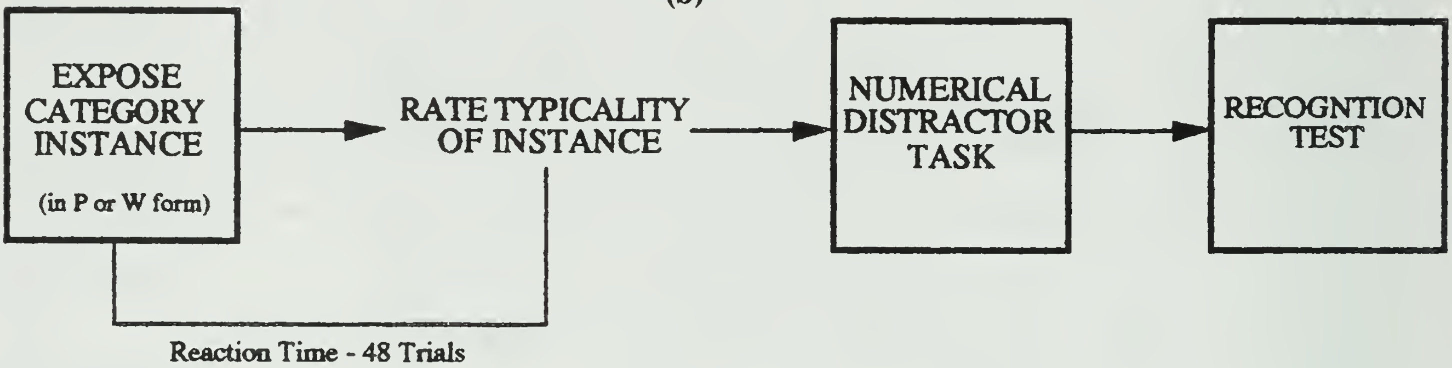
RT for PP = a + b + r
RT for WP = a + b + r
RT for PW = a' + b' = r
RT for WW = a' + b' + r

FIGURE B
OVERVIEW OF EXPERIMENTAL PROCEDURES

Experiment 1
(a)



Experiment 2
(b)



Experiment 3
(c)

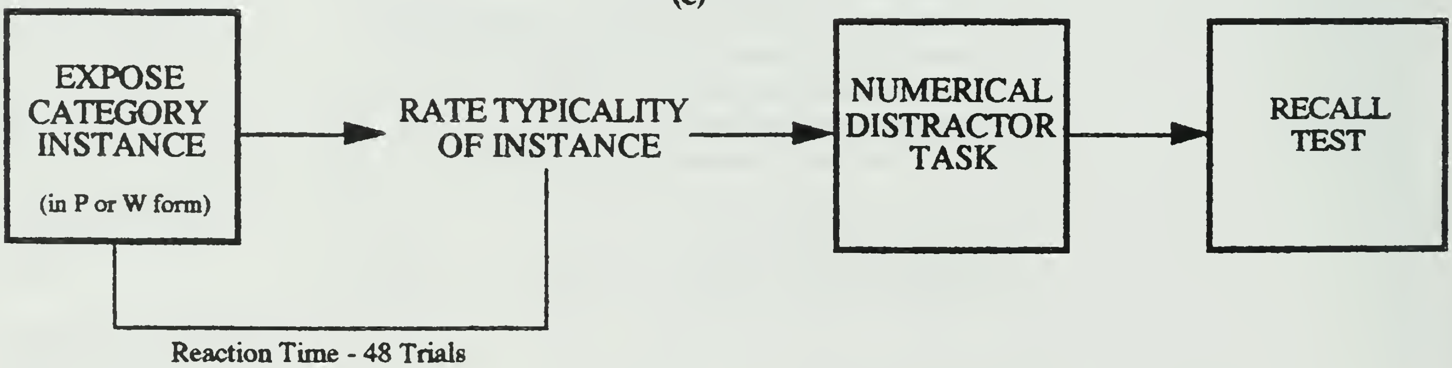
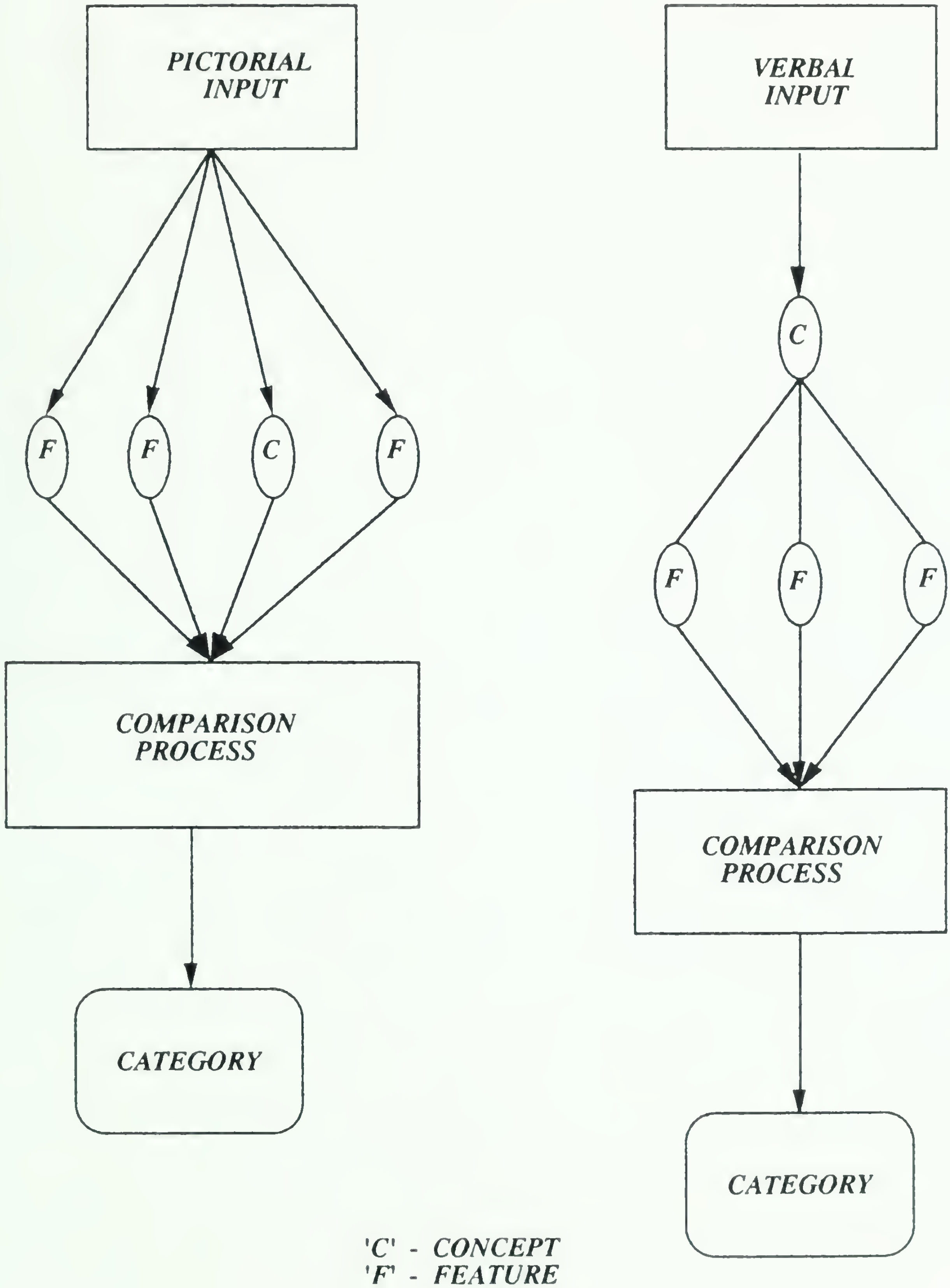


FIGURE C
MODEL OF PICTURE-WORD CATEGORIZATION



'C' - CONCEPT
'F' - FEATURE