THE EFFECTS OF IMAGERY ABILITY, CONCRETENESS AND CONTEXT ON RECALL OF PROSE AMONG BILINGUAL AND MONOLINGUAL SPEAKERS

By

SARAH ELLEN RANSDELL

A DISSERTATION PRESENTED TO THE GRADUATE SCHOOL OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

1987
ACKNOWLEDGEMENTS

There is absolutely no easy way to acknowledge the many contributions of my supervisory committee chairperson, Ira Fischler, to my graduate career. The most tangible, I will try to name. Both academically and emotionally, he intervened only when my own frustrations or limitations made it necessary. He placed me in an environment conducive to growth and always allowed me to make my own choices (and hence errors, at times). I only hope that I have taken something of his skills with me, as I know that a part of him will always influence my life.

I would also like to thank Michael Levy for his constant guidance and leadership these past four years. His knowledge of everything from income tax to computer wizardry has made him indispensable. Richard Griggs, master Jedi of teachers, has forever positively shaped the way I view my life as a teacher of psychology. Walt Cunningham has patiently dealt with all of my woeful inquiries into the world of correlation coefficients. Also, I would like to acknowledge the benefit of my many stimulating conversations with Bill Sullivan.
One person has contributed an invaluable service to my sanity while this dissertation was in progress. My husband John has given me an infallible sense of "I'm behind your work and its importance to you". I will always remember my years in Florida, with him and with my fellow graduate students, as being a period of my life that was both irreplaceable and priceless.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACKNOWLEDGEMENTS</td>
<td></td>
<td>ii</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td></td>
<td>vi</td>
</tr>
<tr>
<td>ABSTRACT</td>
<td></td>
<td>vii</td>
</tr>
<tr>
<td>I</td>
<td>INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>Imagery and Concreteness</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Concreteness and Recall of Prose</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>Components of Imagery Ability</td>
<td>12</td>
<td></td>
</tr>
<tr>
<td>Imagery Ability and Memory</td>
<td>18</td>
<td></td>
</tr>
<tr>
<td>Bilingual Memory and Imagery Strategies</td>
<td>24</td>
<td></td>
</tr>
<tr>
<td>II</td>
<td>EXPERIMENT 1</td>
<td>28</td>
</tr>
<tr>
<td>Method</td>
<td>28</td>
<td></td>
</tr>
<tr>
<td>subjects</td>
<td>28</td>
<td></td>
</tr>
<tr>
<td>materials</td>
<td>28</td>
<td></td>
</tr>
<tr>
<td>procedure</td>
<td>30</td>
<td></td>
</tr>
<tr>
<td>Results</td>
<td>33</td>
<td></td>
</tr>
<tr>
<td>free recall</td>
<td>33</td>
<td></td>
</tr>
<tr>
<td>visual imagery ratings and spatial abilities</td>
<td>37</td>
<td></td>
</tr>
<tr>
<td>summary of results</td>
<td>41</td>
<td></td>
</tr>
<tr>
<td>III</td>
<td>EXPERIMENT 2</td>
<td>44</td>
</tr>
<tr>
<td>Method and Results</td>
<td>45</td>
<td></td>
</tr>
<tr>
<td>Discussion of Experiment 2</td>
<td>47</td>
<td></td>
</tr>
<tr>
<td>IV</td>
<td>EXPERIMENT 3</td>
<td>48</td>
</tr>
<tr>
<td>Method</td>
<td>49</td>
<td></td>
</tr>
<tr>
<td>Results</td>
<td>50</td>
<td></td>
</tr>
<tr>
<td>Discussion of Experiment 3</td>
<td>54</td>
<td></td>
</tr>
<tr>
<td>V</td>
<td>EXPERIMENT 4</td>
<td>55</td>
</tr>
<tr>
<td>Section</td>
<td>Page</td>
<td></td>
</tr>
<tr>
<td>------------------------------------------------------------------------</td>
<td>------</td>
<td></td>
</tr>
<tr>
<td>Method and Results</td>
<td>55</td>
<td></td>
</tr>
<tr>
<td>Discussion of Experiment 4</td>
<td>58</td>
<td></td>
</tr>
<tr>
<td>VI GENERAL DISCUSSION</td>
<td>61</td>
<td></td>
</tr>
<tr>
<td>Manipulating Concreteness Between or Within Subjects</td>
<td>62</td>
<td></td>
</tr>
<tr>
<td>A Question of Retrieval versus Encoding Differences</td>
<td>65</td>
<td></td>
</tr>
<tr>
<td>Effects of Bilingualism on Prose Recall</td>
<td>66</td>
<td></td>
</tr>
<tr>
<td>Visual Imagery Differences: monolinguals</td>
<td>67</td>
<td></td>
</tr>
<tr>
<td>Visual Imagery Differences: bilinguals</td>
<td>68</td>
<td></td>
</tr>
<tr>
<td>Future Research and Extensions</td>
<td>71</td>
<td></td>
</tr>
<tr>
<td>APPENDICES</td>
<td>74</td>
<td></td>
</tr>
<tr>
<td>Appendix A: Four Paragraphs Used as Stimuli</td>
<td>74</td>
<td></td>
</tr>
<tr>
<td>Appendix B: An Example From Each Individual Difference Measure</td>
<td>78</td>
<td></td>
</tr>
<tr>
<td>Appendix C: Language History Questionnaire</td>
<td>79</td>
<td></td>
</tr>
<tr>
<td>REFERENCES</td>
<td>80</td>
<td></td>
</tr>
<tr>
<td>BIOGRAPHICAL SKETCH</td>
<td>84</td>
<td></td>
</tr>
</tbody>
</table>
## LIST OF TABLES

### TABLE

<table>
<thead>
<tr>
<th>Table</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Incidental Free Recall Percent Correct under Comprehension Instructions for Monolinguals versus Bilinguals with Concreteness as a Within-subjects Factor</td>
<td>34</td>
</tr>
<tr>
<td>2</td>
<td>Monolinguals' Means, Standard Deviations, Ranges, and Intercorrelations for Self-reported Visual Imagery, Spatial Ability, and Prose Recall under Comprehension Instructions</td>
<td>38</td>
</tr>
<tr>
<td>3</td>
<td>Bilinguals' Means, Standard Deviations, Ranges, and Intercorrelations for Self-reported Visual Imagery, Spatial Ability, and Prose Recall under Comprehension Instructions</td>
<td>40</td>
</tr>
<tr>
<td>4</td>
<td>Incidental Free Recall Percent Correct under Comprehension Instructions for Monolinguals with Concreteness as a Within versus Between-subjects Factor</td>
<td>45</td>
</tr>
<tr>
<td>5</td>
<td>Incidental Free Recall Percent Correct under Comprehension versus Imagery Rating Instructions for Bilinguals</td>
<td>51</td>
</tr>
<tr>
<td>6</td>
<td>Bilinguals' Means, Standard Deviations, Ranges, and Intercorrelations for Self-reported Visual Imagery, Spatial Ability, and Prose Recall under Imagery Rating Instructions</td>
<td>53</td>
</tr>
<tr>
<td>7</td>
<td>Incidental Free Recall Percent Correct under Imagery Rating Instructions for Monolinguals with Concreteness as a Within versus Between-subjects Factor</td>
<td>57</td>
</tr>
<tr>
<td>8</td>
<td>Monolinguals' Means, Standard Deviations, Ranges, and Intercorrelations for Self-reported Visual Imagery, Spatial Ability, and Prose Recall under Imagery Rating Instructions</td>
<td>58</td>
</tr>
</tbody>
</table>
Abstract of Dissertation Presented to the Graduate School of the University of Florida in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy

THE EFFECTS OF IMAGERY ABILITY, CONCRETENESS AND CONTEXT ON RECALL OF PROSE AMONG BILINGUAL AND MONOLINGUAL SPEAKERS

By

Sarah Ellen Ransdell

August 1987

Chairman: Ira Fischler
Major Department: Psychology

A series of four experiments investigated the interrelationships between recall of abstract and concrete prose passages, individual differences in imagery ability, and strategy differences between monolingual versus bilingual speakers. The first experiment involved a comparison between monolingual and bilingual speakers' incidental free recall of abstract and concrete paragraphs following comprehension instructions. The second experiment contrasted the manipulation of the concreteness of paragraphs as a within- or between-subjects factor. The third and fourth experiments involved changing the instructions that were given preceding incidental free recall. Monolingual and
bilingual subjects now received instructions to rate the imagery value of the sentences that comprised the two types of paragraphs. Included with each of these four experiments' recall data were correlations of the subjects' memory performance with their visual imagery self-ratings and spatial aptitude.

A main finding was that under the comprehension instructions of Experiment 1, the bilingual group did not show a concreteness effect, while the monolingual group showed an almost two-to-one advantage for concrete paragraphs. Native English bilinguals' self-reported visual imagery ratings and spatial aptitude did not appear to correlate highly with their recall of abstract and concrete paragraphs under comprehension instructions; but under the imagery rating conditions, significant correlations were found. Presenting monolinguals with only concrete or abstract paragraphs eliminated the concreteness effect observed in Experiment 1. Both language groups showed a large concreteness effect after the rating tasks (Experiments 3 and 4). It was concluded that the concreteness effect is not due to inherent differences in how concrete and abstract words are represented, but to how a paragraph is processed.
CHAPTER I
INTRODUCTION

Many factors can influence memory for verbal information. Concreteness of the stimuli has proven to be one of the most robust and salient in its effect on memory for word lists, with concrete words typically remembered better than abstract words. The traditional explanation of this advantage is that concrete words have access to perceptual or imaginal types of coding (Paivio, 1971, 1983). However, this concreteness effect has been shown to be more complex and less ubiquitous when looking at memory for connected discourse, suggesting that recall of abstract and concrete paragraphs may be influenced by factors different than those affecting recall of abstract and concrete words.

The present research examined several factors that could possibly account for this apparent dissociation between concreteness effects in memory for prose and isolated words. In addition to varying the concreteness of prose stimuli, individual differences among subjects in visual-spatial imagery ability and preference was assessed and correlated with recall performance. This relationship would give evidence regarding differential
reliance on perception-like coding strategies to remember prose passages. Strategy differences are also possible based on potential access to a second verbal code, as in the case of speakers of two or more languages. The second language of the bilingual may interfere with his or her first, while information stored in a more "imaginal" conceptual code may be less influenced by the two potential verbal codes. There is some evidence that bilinguals may rely to a greater extent on visual imagery strategies than do monolinguals (Ransdell & Fischler, 1987) and therefore this may be an appropriate distinction for uncovering the nature of the concreteness effect.

Imagery and Concreteness

According to work by Paivio, Yuille & Madigan (1968), concreteness can be defined in terms of "directness of reference to sense experience," while imagery can be described in terms of a "word's capacity to arouse nonverbal images." Paivio et al. (1968) determined concreteness, imagery and meaningfulness ratings for 925 nouns and found that about 70 of these words showed reliable differences between concreteness and imagery scores. In particular, words that have relatively high imagery ratings and low concreteness ratings are associated with a sensory experience (usually
affective in nature) but are not tied to specific things or classes of things (e.g., anger, humor, vanity). The specificity of concrete words seems to be the dimension on which concreteness and imagery are most separable. This may provide a clue as to the dissociative effects of concreteness for prose versus words, since concrete and abstract words have inherently more specific referents due to the context of paragraphs.

At the level of word recall (Akin & Schwanenflugel, 1987) and prose recall (Marschark, 1985) there is also some evidence that "concreteness" and "imagery" can be dissociated. One goal of the present research is to uncover the relationship between imagery strategies and concreteness manipulations as they affect prose recall. This chapter will begin with a description of research on prose recall by monolinguals when varying word concreteness. Next, work on individual differences in self-reported visual imagery and spatial ability as correlates of memory will be outlined. Finally, the relationship between bilingual memory and use of concreteness manipulations versus imaginal strategies for recall will be discussed followed by some supporting data.
The main purpose of the research was to detail the impact of concreteness manipulations and imagery strategies on memory for prose, and examine possible strategy differences between speakers of one versus two or more languages. The present research investigated the nature or locus of the concreteness effect, first by looking at its relationship with memory for prose, and second by viewing its relationship with individual differences in visual imagery ability and second language proficiency.

**Concreteness and Recall of Prose**

The relationship between word concreteness and memory has proven to be one of the most consistent effects in cognitive psychology. That concrete, imaginable stimuli are remembered better than more abstract stimuli has been demonstrated in a wide variety of situations (Paivio, 1971, 1983; Yuille & Paivio, 1969). Of more questionable status is the particular locus of this effect. What components of stimuli which elicit high concreteness ratings produce their memory facilitating effects?

Traditionally, an approach based on dual-coding theory (Paivio, 1971) has been used to account for concreteness effects. Stemming from the verbal learning tradition, individual words rated for concreteness have
led to a focus on visual imagery-eliciting qualities in both recognition (Begg & Paivio, 1969) and recall (Paivio, 1971). Paivio's dual-coding theory, developed in conjunction with this type of list-learning research, stated that concrete words obtain their memorial benefits from their ability to access two different "surface" codes, specifically a verbal code and a visual imagery code. This theory has been predominant for a number of years, since it encompassed the results of a large number of verbal learning experiments dealing with the manipulation of word concreteness.

Recently, several researchers have suggested that the concreteness effect is not due to the stimuli's "visual imagery" qualities. Kroll & Merves (1986), for example, have found that when lexical decisions about concrete nouns preceded lexical decisions about abstract nouns, there was a clear effect of concreteness. However, when abstract nouns preceded concrete nouns, the effect disappeared. Since the "imagery" elicited by these nouns is unlikely to change as a function of order of presentation, some other strategic factor must be involved in better performance when concrete nouns are presented first.
Research on prose recall has also produced results that are inconsistent with an "imagery" interpretation of the concreteness effect in memory. It appears that there is no difference in recall of abstract and concrete materials presented in the form of paragraphs. Marschark (1985) did not obtain a concreteness advantage when using prose paragraphs constructed of sentences with concrete versus abstract words, but reinstated the usual concreteness effect when he presented the same sentences in random order. The paragraphs that made up the abstract and concrete version of each paragraph pair were comparable in rated comprehensibility, syntactic structure, and conceptual structure. The main difference between paragraph pairs was the "concreteness" of the key content words within the two paragraph types.

Schwanenflugel & Shoben (1983) presented abstract and concrete paragraphs either with or without a preceding context (three leading sentences) and measured reading time for the target sentence (e.g., "Errors are often made in stressful times."). The concreteness advantage was found only in the "without context" condition, that is, when the sentences were presented alone without the three leading sentences. With a preceding context available, the reading times did not differ. They replicated this finding using lexical decision times with or without a
preceding sentence context. As in Marschark (1985), then, contextual information determined whether a concreteness advantage was obtained. That is, the more contextual constraints placed on the to-be-remembered information, the less impact word concreteness seemed to have. As a final test of this hypothesis, subjects rated the specific context available in a sentence on a scale from 1 (hard to think of a context) to 7 (easy to think of a context) and also the concreteness of the sentence from 1 (low concrete) to 7 (high concrete). Rated context availability and concreteness were found to be highly correlated, $r = .88$. However, Akin & Schwanenflugel (1987) have shown that when words were equated for rated context availability, that is, when the usual confounding of concreteness and context was controlled, recall of abstract and concrete words lists was equivalent. Clearly, a closer look needs to be taken at the parameters involved in the better memory observed for "high imagery" stimuli.

Richardson (1985) tested cued recall and recognition of complex concrete sentences (e.g., "The ants in the kitchen ate the sweet jelly which was on the table") and abstract sentences (e.g., "The intense desire to be successful can determine all personal actions") and found that instructions to use mental imagery had no effect on
either performance measure and that memory was equivalent for the two types of sentences. Context specificity of Richardson's stimuli was similar for both the abstract and concrete sentences, that is, these sentences were probably specific enough to eliminate any potential concreteness advantage. Notice that Richardson's results differ from Marschark's finding of a concreteness advantage with paragraphs when the "complex" sentences within them were presented in random order. It is possible that the sentences in Marschark's study were less context-bound or specific than the longer, more detailed sentences in Richardson's. Clearly, further work in this area is needed since concreteness effects are not universal, especially in studies of memory for connected discourse.

There is additional evidence that contextual availability appears to be more important than word concreteness as a determinant of prose recall. Bransford & Johnson (1975) have shown that under conditions of good context, even low imagery material is remembered well. They presented subjects with a rather abstract paragraph including sentences such as, "First you arrange things into different groups". Those subjects who received knowledge that the topic was "washing clothes" before reading the paragraph recalled more than those subjects
who either received the topic after the paragraph was presented or no topic at all. Conversely, they showed that even high imagery material is retained with difficulty under context-poor conditions. Subjects heard a paragraph that was quite imaginable (first sentence, "If the balloons popped the sound wouldn't be able to carry since everything would be too far away from the correct floor"). Subjects not receiving an initial picture that provided a specific context for the paragraph before reading it performed much more poorly than those who did receive the picture.

Kieras' (1978) context-availability theory can be used to account for findings like Marschark's and others, since memory for connected discourse appears not to be consistently influenced by the concreteness of the stimuli. In Kieras' theory, concrete words are those which can make the contextual situation specific, whereas abstract words produce a more ambiguous interpretation of a particular setting and thus are harder to remember. The context-availability theory could account for the lack of a concreteness advantage, since the relational information in a paragraph may have more influence than the extra amount of context available in concrete words relative to abstract when they are presented in isolation.
Kieras has hypothesized that concrete sentences may be assigned a context more easily than abstract sentences, leading to better performance for the former. Memory for prose is heavily influenced by processing that involves relating the ideas within the text to each other, while less attention is paid to item-specific processing of individual words (Rumelhart, 1977; Sachs, 1967). By increasing the amount of context available in a passage through the inclusion of concrete material, the amount of processing that relates information in a paragraph is also increased (Einstein, McDaniel, Bowers & Stevens, 1984). Therefore, concreteness manipulations should have relatively more impact on stimuli that are not presented within prose, such as words presented in isolation, or have only weak context, as with Marschark's (1985) and Schwanenflugel & Shoben's (1983) isolated or randomized sentences.

Einstein, McDaniel, Bowers & Stevens (1984) have suggested that memory for prose is influenced strongly by the degree of relational (macrolevel) processing versus proposition-specific (microlevel) processing. Microlevel processing is used by subjects when they must focus on word-by-word strategies that make distinctive an individual item in a list. Macrolevel processing, on the other hand, is used to retrieve more thematic, relational
information, such as the gist of the story or paragraph to be remembered. The microlevel, item-specific coding required of subjects in verbal learning tasks such as paired-associate learning or free-recall of word lists, is attenuated with the use of prose material, because macrolevel, relational processing becomes prominent (Kintsch & van Dijk, 1978; Einstein, McDaniel, Bowers, & Stevens, 1984).

There is at least one instance where context specificity and item-specific processing have been controlled for in the same memory experiment. Akin & Schwanenflugel (1987) tested recall of word lists of abstract and concrete nouns that had been equated for context availability and found no effect of concreteness. Since a word recall test presumably involves only item-specific processing rather than relational, then the effects of concreteness cannot be isolated to tasks that require only microlevel coding.

The idea of microlevel and macrolevel processing fits within the general framework of current models of prose comprehension. Comprehension and memory for prose are generally viewed as consisting of several component subprocesses (Kieras, 1981). Current models of the component processes involved in the comprehension of simple prose (Kieras, 1981) and recognition memory
(McGee, 1980) emphasize integrative versus task-specific strategies (i.e. relational versus item-specific processing, respectively) at encoding rather than the imaginable properties of the stimuli. Therefore, evidence of concreteness by context interactions must be viewed in terms of the type of processing required of the subject and also in terms of the amount of context available in the stimuli.

Components of Imagery Ability

Another approach to examining the effects of concreteness involves the investigation of the relationship between individual differences in imagery ability and memory performance. Viewing individual differences in imagery abilities may help clarify the nature of concreteness effects. Evidence for a greater concreteness advantage among "high-ability" subjects than among "low-ability" subjects would imply that concreteness effects in prose recall are at least partly imagery-linked.

Imagery ability has typically been measured through the use of self-report questionnaires, on the one hand, and spatial abilities tests and other performance-based measures on the other. Imagery ability has been shown to predict performance in situations where the availability
of imagery as a coding device or a processing strategy is manipulated.

The search for different components of imagery and imagery ability arose out of the inadequacy of depicting imagery as a "monolithic" concept (Kosslyn, Brunn, Cave & Wallach, 1984). Kosslyn et al. gave subjects a series of imagery tasks and looked for evidence of distinct, independent "skills", that is, essentially zero correlations among sets of scores. The logic behind this analysis implies that the correlation between any two tasks will depend on the number of shared processing components, with higher numbers shared being reflected by higher correlations. A very wide range of correlation coefficients were found, suggesting that the subjects were not simply good or poor at imagery in general, but could be differentiated on a subset of relatively distinct imagery abilities.

One major distinction in the measurement of imagery ability lies in the use of objective performance tests versus self-report questionnaires. Poltrock & Brown (1984) find evidence that spatial performance tests and self-reports of imagery ability apparently measure different aspects of imagery. They claim that the coarse nature of self-report ratings prevents such measures from serving as a useful way to develop a theory of imagery.
Although a moderate relationship was found to exist between rated imagery control and a spatial test, other possible correlations between self-ratings and performance were lacking. Although people may be aware, at least to some degree, of their cognitive strengths and weaknesses, the validity of verbal reports as data has been questioned (Ericsson & Simon, 1980). In particular, subjective reports are more likely to provide an accurate description of actual processes when 1) the processes are not very demanding, but at the same time have not become automatized, and 2) the reports concern concurrent, rather than prior events. The self-report questionnaires typically used in visual imagery studies appear to have these characteristics. Also, some proposed components of imagery ability, such as vividness (see below), stress phenomenological aspects of imagery that are properly measured by self-reports.

The most likely salient components of imagery ability have been explored in several studies (see below). For the purposes of the present study, those tests that capture the variance of these components need to be selected. White, Sheehan & Ashton (1977) have surveyed the current self-report measures of individual differences in imagery ability and have made several recommendations. Based on the psychometric properties of
these self-report measures, White et al. suggested that the reliability and validity of several tests be considered adequate for use in correlational studies. They also suggest that what imagery is defined as and what its function will be in a given task should be distinguished to determine precisely what a given test item measures. An additional important consideration concerns expectancy factors, such as social desirability and test response sets.

Lorenz & Neisser (1985) looked at the relationship between nine different mental imagery measures and childhood memory and found evidence for three factorially distinct imagery factors, "vividness and control", "spatial manipulation", and "spontaneous elaboration". Bett's QMI, Marks' VVIQ, and Gordon's TVIC were seen to measure vividness and control. The differential aptitude test of space relations and a cube-cutting task measured spatial manipulation, and Barrett's visualization of forms and Richardson's VVQ were measures of spontaneous elaboration. Other researchers have found a similar pattern roughly corresponding to these three factors, although vividness and control have been found to have separable characteristics (Richardson, 1977a) and spontaneous elaboration has more usually been referred to as a cognitive style dimension (Richardson, 1978). It is
important to note that most of these test measures were developed in conjunction with some type of memory performance correlates in mind. For instance, Marks (1973) found evidence that subjects who reported vivid visual imagery (VVIQ) were more accurate in picture recall than subjects who reported poor visual imagery.

Poltrock & Brown (1984) found the VVIQ to be an adequate measure of a factor they identified as image quality, and the TVIC as indicative of image process efficiency. Subjects participated in six laboratory tasks that were devised to measure image quality and the efficiency of image generation, image rotation, image scanning, adding, and subtracting detail in images, and integration of images. The laboratory tasks correlated with the two self-report measures, VVIQ and TVIC, according to whether the tasks were considered to tap predominately image quality versus image processing efficiency skills, respectively. Selection of tests of both vividness and controllability for the present research were thus influenced by the above findings, especially since they were found to be factorially distinct and separable under some conditions. Imagery vividness has also often been measured by the Betts QMI (Richardson, 1969) in addition to the Vividness of Visual Imagery Questionnaire, VVIQ (Marks, 1973). The VVIQ was
selected since it focuses on the visual modality (the QMI covers seven sensory modalities) and because evidence for relatively strong correlations with social desirability (r=.46) have been shown for the QMI (Durndell & Wetherick, 1975).

By far the best known subjective test for imagery control has been the Gordon Test of Visual Imagery Control, TVIC (Gordon, 1949). Since the TVIC appears to have high reliability (internal .64 to .95, and test-retest .60 to .84) it has been chosen to measure this ability.

A third major dimension concerns the verbalizer-visualizer dichotomy. Paivio's Individual Differences Questionnaire (IDQ) was designed to assess habitual thinking modes as either imaginal or verbal in nature. In an extension of Paivio's work, Richardson (1977b) developed a shorter test which he calls the Verbalizer-Visualizer Questionnaire (VQ). This latter test has been selected as its reliability is high (r=.91 to .93) and it has a very low correlation with social desirability (r=.02). Spatial ability can be measured by many different tests (all of which appear to be highly correlated) but a form of the Differential Aptitudes Test, space relations, has been selected since it is a
widely used test in factor analytic studies of spatial aptitude.

**Imagery Ability and Memory**

If individual differences in the imagery factors of vividness, control, preference and/or spatial ability correlate with evidence of a concreteness advantage, with high "imagers" producing higher levels of recall for concrete prose, then phenomenological reports of visual imagery as a memory strategy behind the concreteness effect would be given some credence. However, if no reliable correlations are found between concreteness effects in memory for prose and imagery ability, then the possibility that these tests do not tap imaginal abilities and strategies must be considered. In the following section, several results are presented to lend support to the idea that self-reported imagery ability does in fact correlate with memory performance. Given evidence for imagery ability correlates with memory, it can be shown that "visual imagery", as measured by self-reports, or by performance on spatial abilities tests, is not the basis of differential memory for abstract and concrete stimuli when equivalent contextual information is available to both concrete and abstract stimuli.
Individual differences in visual imagery vividness, controllability, spatial aptitude and coding preference have been proposed as factors that predict differences in memory for abstract and concrete materials (see Ernest, 1977 for a review). Two self-report measures, vividness and control, have been shown to correlate with memory performance in incidental tasks. And for tasks such as mental rotation, paired associate learning, problem solving and other spatial tasks (Poltrock & Brown, 1984) correlations have been found with self-rated controllability.

In evaluating any correlations between self-rated vividness of visual imagery and subsequent recall performance, it is important to distinguish between incidental and intentional recall. Image quality or vividness appears to have a relatively low correlation with intentional learning of word lists. Sheehan (1973) found an increase in the correlation between the effects of word concreteness and imagery ability in incidental memory situations. It appears that the ability to experience vivid imagery facilitates retrieval of detail or content information from memory, but has less impact when the task requires the explicit use of strategies for recall of connected discourse. Vivid imagery is used as a processing strategy by those who score "high" on this
dimension more often if the experiment encourages, cues, or primes its use. In the case of recall from prose, higher order relational strategies may be called for, thereby eliminating the need for any micro-level strategies based on the visual imagery qualities of specific words.

In typical list-learning tasks, where the emphasis is more on item-specific strategies (those strategies that focus on the imagery value of individual words), differences found between high and low imagers are at their strongest under abstract or low imagery conditions. That is, high imagery ability helps the most with low imagery stimuli, particularly when subjects must pay attention to word by word information, (Ernest, 1977). This can be contrasted with memory for prose where high imagery ability is usually associated with overall superior performance, but where interactions reveal that low imagers either surpass or equal high imagers when the passage is abstract.

For example, Riding & Taylor (1976) developed two performance tests of imagery ability and searched for correlations with memory for concrete and abstract prose material. An image-generation test was based on the speed with which an image of an inverted object could be generated. A memory-code test measured the speed with
which subjects could generate new information about objects or incidents described in a short story. In a study with children, Riding & Taylor found better overall memory for prose passages for children of high imagery ability as measured by both objective tests, but found that the low imagers actually did better on the abstract passages. They concluded that meaningful context encourages image arousal in high imagers, but not in low imagers. High imagers are particularly good at maintaining in memory a "literal" representation of stimuli so that memory for more abstract prose is not always aided by this skill.

It is interesting that individual differences in imagery ability seem to have different relationships with memory for word lists versus prose. The above research points to the conclusion that high imagers remember concrete words, abstract words, and concrete prose better than low imagers, but that low imagers may have the edge in remembering abstract prose. Hiscock (1976) showed that visualizers outperformed verbalizers in recalling high-imagery adjectives within prose, but did slightly worse than verbalizers when adjectives were low-imagery. Other researchers (Janssen, 1976; Giesen & Peeck, 1984) have not replicated this result, since their measures of imagery did not yield a significant correlation with
performance. Further research is needed in this area since with some measures of imagery ability, high-imagers appear most successful when material doesn't spontaneously elicit imagery, while other work suggests that with abstract material, high-imagers will be at a disadvantage. Research that attempts to reconcile these apparently inconsistent findings will need to include more than a single dimension of visual imagery ability and will need to look at memory for both word lists and for prose.

Imagery control, compared to measures of vividness, appears to have, in general, higher correlations with memory for concrete versus abstract material. The ability to manipulate and control a visual image, regardless of its clarity, parallels spatial tasks as a predictor of associative learning, serial recall and picture/word list differences (Lansman, Donaldson, Hunt & Yantis, 1982). Subjects high in spatial control excelled in particular when generating images to abstract words, comprehending abstract sentences, recognizing unfamiliar pictures, and learning material in the absence of instructions to learn (Ernest, 1977). As with vividness, control and spatial-imagery ability are most facilitative with low imagery stimuli, but again complete interpretation is not possible without considering scores
on more than one imagery dimension, and the nature of the memory test.

Coding preference has also been used to assess individual differences in memory performance. Richardson (1977b) developed a test for measuring the likelihood that a subject will encode a verbal stimulus as a mental image. Habitual use of imagery has been shown to be predictive of some aspects of verbal recall (Poltrock & Brown, 1984). Self-reports of being a visualizer on the VVQ test were found to be correlated with the efficiency with which an individual can add image detail and mentally rotate images.

All of these self-report measures seem to have some relationship with memory performance for abstract and concrete stimuli, especially when the experimental environment is less than maximally conducive to image arousal. Memory for abstract prose may be remembered better by low-imagers, while concrete prose and both types of words may be better recalled by high-imagers. The answer then to this interaction lies in determining both the types of individual difference measures of ability, and in the type of memory task. The present research will include measures of vividness, control, preference, and spatial aptitude in order to determine these abilities influence on memory performance for
abstract and concrete paragraphs. Strong correlations between individual differences in imagery ability and free recall will imply that even at the level of prose, visual imagery ability may influence memory. On the other hand, it also may prove to be the case that when some higher order processing can take precedence over imagery strategies, as in the case of connected discourse (Marschark, 1985), imagery ability will have little or no impact on recall.

**Bilingual Memory and Imagery Strategies**

Differences in second language ability have also been shown to influence imagery strategies. Magiste (1980) found that bilinguals showed higher error rates in recall of two-digit numbers and in object-naming compared to monolinguals. She suggested that the bilingual disadvantage could be due to differences in familiarity with the target language, or to a more specific interference of the second language on performance of the first. The familiarity with a target language (e.g. English) can be made more comparable by using native English bilinguals who are also fluent in a second language. Ransdell & Fischler (in press) reported such an experiment with native English bilinguals and found an interaction of language (monolingual versus bilingual subjects) and memory for abstract versus concrete nouns.
The bilinguals were as fast as the monolinguals to recognize concrete words, while they were substantially slower in the abstract condition. Bilinguals, then, showed an effect of concreteness when monolinguals did not. Paivio & Desrochers (1980) hypothesized that any difference between bilinguals and monolinguals would be exaggerated for concrete words, since these would allow for "cross-talk" between languages via the imagery system.

Having a second verbal code might influence the relationship or balance of verbal and visual imagery strategy use. There is evidence that "true" bilinguals, those who alternate frequently between their two languages in everyday life, tend to disregard the verbal labels attached to various visual scenes, even when this does not contribute to solving a problem efficiently (Miljkovitch, 1980). Miljkovitch found that bilinguals group words from two languages into categories that receive no verbal label in either language. One important exception to this nonverbal strategy may occur when relational processing inherent in prose passages makes the use of word by word imagery mnemonics less beneficial.
If individual differences in "imagery" ability do not significantly correlate with memory performance for abstract and concrete prose material then this will suggest that visual imagery ability is not an important factor behind recall of information at the level of connected discourse. Since for monolinguals the shift from words in isolation to prose drastically changes the pattern of concreteness effects, the contrast between the two language groups might best be studied with prose. Of special interest is whether bilingual subjects will make more use of imaginal strategies and show more visual imagery aptitude on self-report measures relative to monolinguals, as predicted by Paivio & Desrochers. It could be the case that bilinguals are more visually oriented because of the linguistic code ambiguity produced by their two languages. Alternatively, bilinguals may be less "visual", since their second language may create greater sensitivity to linguistic distinctions, although the above research does not suggest this.

The first experiment involved a comparison of monolingual and bilingual speakers' incidental free recall of abstract and concrete paragraphs following comprehension instructions. Included with the recall data are correlations of the subjects' memory performance
with their visual imagery self-ratings and spatial aptitude. The specific predictions are as follows: 1) overall, bilinguals may show an exaggerated concreteness advantage even with prose, since they may rely on visual strategies more so than monolinguals; 2) monolinguals will be less likely to show a concreteness advantage since previous research shows that context-availability has a stronger effect on prose memory than does word concreteness; 3) bilinguals will show higher absolute scores on the imagery self-reports and/or spatial ability measures; 4) if imagery ability has anything to do with concreteness and prose recall, bilinguals will show greater correlations of these individual difference measures with prose recall performance than monolinguals; and 5) those subjects who score above the median on measures of vividness and control will perform better on incidental recall of concrete prose, while low scorers will perform better on recall of the abstract.
Subjects

Forty-eight University of Florida undergraduates, 22 women and 26 men, participated for partial fulfillment of a course requirement. Subjects were randomly assigned to one of 16 possible orders of receiving two abstract and two concrete paragraphs. An additional group of 48 native English bilinguals from the same subject population, 15 women and 33 men, were also in the experiment. Native English bilinguals were defined as those individuals who learned both of their languages as young children, with English learned first or at the same time as a second language. These subjects did not know that they were chosen for the present set of experiments because they were bilingual.

Materials

Four pairs of high- and low-imagery English paragraphs were drawn from those used by Marschark (1985). The paragraphs in each of the four pairs were
equated for number of sentences (5, 5, 6, and 6) and number of words per sentence (means = 15, 16, 15, and 14). Pair members also had identical grammatical and conceptual structures (i.e., story line of the paragraph pairs were the same) as well as predominantly identical wording; they differed essentially only in the imageability of their key content words. The mean imageability rating of the sentences was 5.18 for the high-imagery paragraphs and 3.09 for the low-imagery paragraphs. In addition to the four paragraphs used in experiment 1, one paragraph of each type was also taken from Marschark to be used as a practice paragraph. Samples of each type of paragraph are given in Appendix A. One paragraph was printed on each page and two paragraphs of each type were compiled into test booklets according to a partial Latin square design, counterbalancing for order and paragraph type.

Subjects also received three of the self-report imagery questionnaires described above and a spatial aptitude test. Samples of material from each of the tests are given in Appendix B. Mark's (1973) Vividness of Visual Imagery Questionnaire (VVIQ) assessed visual quality or vividness, Gordon's (1949) Test of Visual Imagery Control (TVIC) served as a measure of subject's subjective imagery controllability, Richardson's (1977b)
Verbalizer-Visualizer Preference Test (VVQ) assessed preference for strategies along a verbal versus visual dimension. The Differential Space Relations Test (SRA for accuracy and SRT for latency) was used as a performance measure of spatial aptitude (Educational Testing Service, 1972).

Procedure

The procedure described followed as closely as possible Marschark's (1985) Experiment 1. Any differences that exist between the two studies will be pointed out. It is important to note subjects' expectations upon entering the experimental situation. The purpose of the experiment was described as an investigation of comprehension of prose. Subjects did not know that they would subsequently be asked for free recall, nor did they anticipate any imagery ability measures after the paragraph recall.

All subjects, in groups of between 2 and 15, were seated at desks or tables and given the following instructions. Subjects were told that the experiment involved comprehension of paragraphs and would last approximately one hour. They were also told that all responses would remain anonymous and that they should work as quickly and as accurately as they could. The
following information was given to all subjects after paragraph booklets and blank paper were passed out.

"On your desk, face down, is a booklet containing four paragraphs. Your job will be to read each, one at a time, after which questions will be asked about their content. I will tell you when to start and you will have one minute to carefully read each one. Let's look at a single paragraph just for practice. Turn over the top page, which is a single sheet, and look it over for one minute to get an idea of how much time you actually have for each one. Now let's begin the experiment. I will tell you when to go on to the next paragraph. Begin with paragraph one."

After subjects read the last paragraph, booklets were collected and subjects put their respective booklet number on the blank sheet of paper to designate which order and type of paragraphs they received. Next, subjects were told to try to recall as much as possible in any order from the four paragraphs, using the original wording whenever possible. They were told that if they couldn't remember the original wording, to write whatever they could, sentence parts or gist. Subjects were given thirteen minutes for recall of the four paragraphs, which was the time allocated for the subjects in Marschark's
(1985) experiment. At this point, subjects' recalls were collected and they were told that they would now be filling out some brief questionnaires concerning three separate aspects of the use of visual mental imagery.

Subjects were given approximately five minutes to fill out the VVIQ (16 items). Then they were given the TVIC (12 items) and had 3 minutes to complete it. It was stressed that this second questionnaire was assessing a separate aspect of imagery from vividness. The third questionnaire, the VVQ (15 items) was given next, with 4 minutes allowed for completion. Finally, subjects were given the Space Relations test. Both accuracy and speed were measured for each individual. Subjects had up to 14 minutes to complete 30 items and were instructed to raise their hand when they finished so completion time could be noted.

A brief handedness questionnaire was given last and the monolingual subjects were finished with the experiment at this point. Bilingual subjects were then given a Language History Questionnaire to ascertain their usage pattern of each language over time. Part of this questionnaire involved self-ratings of written and spoken linguistic ability in the bilingual's two languages, with a scale ranging from 1-very fluent to 5-not fluent. English, the native language for these speakers, was
rated 1.35 out of 5 for written fluency and 1.16 for spoken fluency. The bilingual's second language was rated 2.54 for written and 2.31 for spoken fluency, suggesting that these subjects were dominant in English, but were relatively fluent in their second language. The language history questionnaire is presented in Appendix C.

Results

Free Recall

All recall protocols were scored independently by the experimenter and an undergraduate assistant naive as to the experimental predictions but trained in scoring prose materials according to gist criteria. Interrater reliability was high as there was almost complete agreement in the scores obtained by the two judges. Out of 192 possible comparisons of data points (48 subjects by 4 scores for the monolinguals in Experiment 1) only 17 scores varied by more than one idea unit. Ten of the 17 scores were more conservative for one scorer and 7, more liberal. Thirteen scores out of 192 varied by more than one idea unit for the bilinguals in Experiment 1, 12 more conservative and 1 more liberal. Due to the complexity of the materials (see Marschark, 1985), idea units were based on propositional structure, but this is not a possible confound since the units for concrete versus
abstract paragraphs were the same. All effects referred to as reliable exceed the .05 level of confidence.

Incidental free recall percent correct for the bilinguals and monolinguals for concrete and abstract paragraphs is presented in Table 1.

Table 1
Incidental Free Recall Percent Correct under Comprehension Instructions for Monolinguals versus Bilinguals With Concreteness Within-subjects

<table>
<thead>
<tr>
<th></th>
<th>Abstract</th>
<th>Concrete</th>
<th>Difference</th>
<th>Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Monolinguals</td>
<td>41.9</td>
<td>70.8</td>
<td>28.9</td>
<td>56.4</td>
</tr>
<tr>
<td>Bilinguals</td>
<td>51.3</td>
<td>57.9</td>
<td>6.6</td>
<td>54.6</td>
</tr>
</tbody>
</table>

In the overall two-way analysis of variance treating language as a between-groups factor and concreteness as a within-subjects factor, there was no significant difference in recall between bilinguals and monolinguals, $F(1,94) = 1.49, p > .10$. An advantage of concrete materials over abstract was found, $F(1,94) = 38.09, p < .001$, and there was an interaction between language and concreteness, $F(1,94) = 8.01, p < .005$. The monolinguals showed a significant 28.9% recall advantage for concrete paragraphs over abstract (70.8% - 41.9%, respectively). For the bilinguals, the 6.6% concreteness advantage (57.9% - 51.3%) was not significant.
For the monolingual group, the most basic finding based on proportion of idea units recalled was a large effect of concreteness. Mean proportion recall for the two abstract paragraphs given to each subject was 41.9% versus 70.8% for the concrete paragraphs. In a one-way analysis of variance treating concreteness as a within-subjects factor, this was a significant effect, \( F(1,47) = 38.63, p < .001 \). In a two-way ANOVA treating order as a between-subjects factor, order of presentation failed to reach significance, as did the interaction between order and concreteness, both \( F \)'s < 1. Neither the effect of practice paragraph (abstract or concrete), \( F(1,47) = 2.70 \), nor the interaction between concreteness and practice were significant, \( F < 1 \).

An ANOVA treating items (that is, particular paragraph) as the random effect was also performed. The difference between abstract and concrete paragraphs was still significant, \( F(15,47) = 84.93, p < .001 \), while the main effect of paragraph was not, \( F = 1.49, p < .10 \). The interaction between paragraph and concreteness was significant, \( F(15,47) = 8.98, p < .001 \). In an analysis of variance treating subjects as a random effect, concreteness remained a significant effect, \( F(1,188) = 78.37, p < .001 \), but now the effect of particular paragraph was also significant, \( F(3,188) = 14.65, p < .001 \). As might be suspected, the interaction between
these two variables continued to be significant, $F(3,376) = 6.60, p < .001$.

In contrast to the monolingual subjects, the mean proportion recall for the bilinguals was 51.3% for abstract paragraphs, and 57.9% for concrete, a difference of 6.6%, compared to a 28.9% difference for the monolinguals. In an analysis of variance with concreteness as a within-subjects factor and items as the random effect, the concreteness advantage was not significant, $F(1,88) = 1.26, p < .10$. As with the same analysis for the monolingual group, the effect of paragraph was significant, $F(3,88) = 3.00, p < .05$. The interaction was not, $F(1,88) < 1$.

The main effect of order was not significant, $F < 1$, nor was the interaction of order and concreteness, $F < 1$.

Since there were more incidences of zero recalls for the abstract versus the concrete paragraphs, (cases where subjects could not remember any of the contents of a given paragraph), proportion correct scores was adjusted for total number of recall attempts to look at the effect of retrieval on recall. When computing mean proportion correct adjusting for the total number of recalls for a given paragraph type, 63.98% of the abstract paragraphs were recalled compared to 75.5% of the concrete paragraphs (41.9% versus 70.8% before adjustment). Since
a difference of 11.52% between the concrete and abstract paragraphs was found when adjusting for zero recalls, then at least some of the discrepancy between abstract and concrete recall can be localized to events that occur at retrieval.

**Visual Imagery ratings and Spatial Abilities**

Visual imagery self-reported ratings and spatial ability (accuracy and latency) were scored in standard procedure for each subject. For the VVIQ, TVIC, and VVQ, responses could be from 1 (highest ability) to 5 (lowest ability). The totals from all items within each of the three self-rating questionnaires were added and then converted to mean scores out of the response range from 1 to 5. Also, rather than interpreting scores by noting that the lower the total, the higher that particular ability, the signs of the correlations with these three measures and other performance tests was reversed so that higher rated ability reflects a higher score. The space relations test, SRA (total percent correct) and SRT (total time taken to complete the 30 items) was scored for each subject.

Correlations among the self-report measures, spatial ability, and recall for abstract and concrete paragraphs for monolinguals is presented in Table 2, along with means, standard deviations and ranges.
Table 2
Monolinguals' Means, Standard Deviations, Ranges, and Intercorrelations for Self-reported Visual-imagery, Spatial Ability and Prose Recall Under Comprehension Instructions

<table>
<thead>
<tr>
<th></th>
<th>VVIQ</th>
<th>TVIC</th>
<th>VVQ</th>
<th>SRA</th>
<th>SRT</th>
<th>ABST</th>
<th>CONC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>1.93</td>
<td>1.80</td>
<td>2.87</td>
<td>18.8</td>
<td>12.9</td>
<td>41.9</td>
<td>70.8</td>
</tr>
<tr>
<td>Std Dev</td>
<td>.45</td>
<td>.40</td>
<td>.30</td>
<td>5.5</td>
<td>2.3</td>
<td>14.1</td>
<td>18.7</td>
</tr>
<tr>
<td>Range</td>
<td>1-3</td>
<td>1-4</td>
<td>2-4</td>
<td>4-29</td>
<td>8-19</td>
<td>0-100</td>
<td>0-100</td>
</tr>
</tbody>
</table>

Intercorrelations

<table>
<thead>
<tr>
<th></th>
<th>VVIQ</th>
<th>TVIC</th>
<th>VVQ</th>
<th>SRA</th>
<th>SRT</th>
<th>ABST</th>
<th>CONC</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>VVIQ</td>
<td>---</td>
<td>.44**</td>
<td>.34*</td>
<td>.23</td>
<td>.17</td>
<td>.08</td>
<td>.16</td>
<td></td>
</tr>
<tr>
<td>TVIC</td>
<td></td>
<td>---</td>
<td>.36*</td>
<td>.08</td>
<td>.16</td>
<td>.10</td>
<td>.14</td>
<td></td>
</tr>
<tr>
<td>VVQ</td>
<td></td>
<td></td>
<td>---</td>
<td>.05</td>
<td>.11</td>
<td>.16</td>
<td>.00</td>
<td></td>
</tr>
<tr>
<td>SRA</td>
<td></td>
<td></td>
<td></td>
<td>---</td>
<td>-.03</td>
<td>.30*</td>
<td>.35*</td>
<td></td>
</tr>
<tr>
<td>SRT</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>---</td>
<td>.14</td>
<td>.17</td>
<td></td>
</tr>
</tbody>
</table>

* p < .05, ** p < .001

Note. Scores for VVIQ, TVIC, and VVQ are based on 16, 12 and 15 items, respectively, with possible responses between 1-high ability to 5-low ability. Therefore, signs have been reflected on all correlations with these three variables since for all other variables, a higher score implies better performance.

Among the "subjective" self-report measures, significant correlations were found between the VVIQ (vividness) and the TVIC (control), r = +.44; the VVIQ and the VVQ (preference), r = +.34; the TVIC and the VVQ, r = +.36. For the performance measures of the space relations test and free recall performance, two correlations were found to be reliable. It is important to note that the correlation between SRA and SRT was only...
$r = +.11$, indicating that subjects did not trade off accuracy for speed on the spatial performance test. SRA (space relations accuracy) correlated significantly with both abstract recall, $r = +.30$, and with concrete recall, $r = +.35$. There were no significant correlations between the self-ratings of imagery ability and any of the performance measures.

A correlation matrix among all individual imagery measures and recall performance for bilinguals is presented in Table 3. The only correlation that was significant among the individual difference measures existed between the VVIQ and the TVIC $r = +.47$. Among the performance measures of recall and spatial ability, a marginally significant correlation was found between SRA (accuracy) and concrete recall, $r = +.41$, $p < .06$. 
Table 3
Bilinguals' Means, Standard Deviations, Ranges, and Intercorrelations for Self-reported Visual-Imagery, Spatial Ability, and Prose Recall Under Comprehension Instructions

<table>
<thead>
<tr>
<th></th>
<th>VVIQ</th>
<th>TVIC</th>
<th>VVQ</th>
<th>SRA</th>
<th>SRT</th>
<th>ABST</th>
<th>CONC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>2.11</td>
<td>1.90</td>
<td>2.84</td>
<td>19.1</td>
<td>12.2</td>
<td>51.3</td>
<td>57.9</td>
</tr>
<tr>
<td>Std Dev</td>
<td>.60</td>
<td>.38</td>
<td>.27</td>
<td>5.5</td>
<td>1.8</td>
<td>22.2</td>
<td>19.2</td>
</tr>
<tr>
<td>Range</td>
<td>1-4</td>
<td>1-3</td>
<td>2-4</td>
<td>8-29</td>
<td>7-15</td>
<td>0-100</td>
<td>0-100</td>
</tr>
</tbody>
</table>

Intercorrelations

<table>
<thead>
<tr>
<th></th>
<th>VVIQ</th>
<th>---</th>
<th>.47**</th>
<th>.04</th>
<th>.09</th>
<th>.24</th>
<th>.08</th>
<th>.16</th>
</tr>
</thead>
<tbody>
<tr>
<td>TVIC</td>
<td>---</td>
<td>.15</td>
<td>.13</td>
<td>.04</td>
<td>.05</td>
<td>.35</td>
<td></td>
<td></td>
</tr>
<tr>
<td>VVQ</td>
<td>---</td>
<td>.05</td>
<td>.04</td>
<td>.17</td>
<td>.07</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SRA</td>
<td>---</td>
<td></td>
<td>-.09</td>
<td>.14</td>
<td>.41*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SRT</td>
<td>---</td>
<td></td>
<td></td>
<td>-.06</td>
<td>-.28</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*p<.06, **p<.001

In order to assess the differences in subjective visual imagery ratings and spatial aptitude between monolingual and bilingual subjects, a multiple analysis of variance treating language as a between-subjects factor and the individual difference measures as dependent variables was performed. There were no significant differences between the two language groups, either on how they rated their imagery ability (VVIQ, F(1,95) = 2.63; TVIC and VVQ, both F's < 1), or on accuracy (SRA, F < 1) or latency (SRT, F = 2.04) of the space relations test.
Summary of Results

The main findings of Experiment 1 are the following: 1) overall recall was equivalent for the two language groups; 2) monolinguals showed a concreteness advantage and bilinguals did not; 3) no significant difference was found in the way the two language groups rated their imagery ability or performed on a spatial task; 4) for both language groups, relatively little contribution of imagery ratings to recall performance was found; and 5) reliable correlations between space relations accuracy and recall of concrete paragraphs was found in both language groups, and in the monolingual group, also with abstract recall performance.

The pattern of results for both the monolingual and bilingual groups were the opposite of what was predicted. Prose recall of abstract and concrete paragraphs was hypothesized to be equivalent, since the amount of context inherent in connected discourse should have produced more use of relational processing strategies relative to item-specific that tend to favor concrete stimuli. Even though monolinguals and bilinguals performed similarly on the visual-spatial imagery measures, the latter did not show a memorial advantage for concrete stimuli. One way to understand the effects of concreteness in the above results would be to
experimentally produce or eliminate its appearance. Therefore, in Experiment 2, an attempt was made to eliminate the concreteness advantage in monolingual subjects, while in Experiment 3, bilingual subjects are asked to focus on the imaginal qualities of the materials in order to promote a concreteness effect.

Since in Experiment 1, monolingual prose recall was influenced by concreteness, while in Marschark (1985) it was not, several possible differences in the two studies were considered before selecting the factor of within-versus between-subject concreteness as critical. First, overall level of recall was lower in Marschark's study, and specific paragraphs of Experiment 1 were not identical to his, although drawn from the same sample. Although overall recall was better in the present study, there was no obvious reason why the concreteness effect would vary systematically with overall recall level. Second, different scoring criteria, that is, more or less conservative, may have been adopted in the two studies. But nominal scoring criteria were identical in the two studies as assessed by personal communication (Marschark, 1987).

Marschark's (1985) finding of the absence of concreteness effects in memory for prose was replicated in an unpublished study by Bunn, which is presented in
Hunt and Marschark (1987). As in the present study, Bunn varied concreteness as a within-subjects variable, but tested recall after each paragraph, and also found a concreteness advantage. When she considered data from the recall of the first presentation for all subjects (comparable to a between-subjects comparison) no concreteness effect was obtained. Hunt & Marschark (1987) explain this finding by focusing on the type of processing that subjects use during the first presentation ("thematic" or shared information typical of prose processing) and in subsequent presentations after the first recall test where memory demands avert the subject's attention to "distinctive" unit-by-unit encoding strategies. Since they argue that distinctive item-specific processing favors concrete materials, concrete paragraphs are recalled more often when subjects receive more than one presentation followed by a memory test. Note that this argument is 1) ad hoc, 2) an encoding, as opposed to a retrieval explanation, and 3) perhaps inconsistent with the fact that "memory demands" of gist recall of prose should be heavily dependent on concept-driven processes, and this type of task is likely to bias relational processing, rather than distinctive.
CHAPTER III
EXPERIMENT 2

It was of interest, based on Hunt & Marschark's interpretation of Bunn's data, to isolate recall of the first presented paragraph in Experiment 1. A mixed 2 x 2 analysis of variance was performed on these data, with concreteness as a between-groups factor and sequence as as a within-subjects factor. Concreteness was still a significant effect for recall of the first presentation paragraph type, $F = 52.36, p < .0001$, while sequence (first, second, third or fourth) was not significant, $F = 2.91, p < .10$, nor was the interaction, $F < 1$.

The above results do not lend themselves well to the preceding discussion of Bunn's unpublished findings. The concreteness effect for the first paragraph is inconsistent with Bunn's results, but the comparison is not completely appropriate because of the timing of recall. Subjects in experiment 1 recalled material from all four paragraphs (two concrete, two abstract) after presentation of the entire set. Since there was not an intervening recall test after each presentation in the present Experiment 1, the result of analyzing the initial presentation cannot be considered a definitive test of
the hypothesis concerning within- versus between-subject manipulation of concreteness. Therefore, Experiment 2 (between-subjects concreteness) was needed as an essential manipulation with a recall attempt after all paragraphs had been presented.

Method and Results

Forty monolingual subjects from the same subject population (University of Florida introductory psychology students) received the identical procedure as in Experiment 1, but received paragraphs of only one type of concreteness.

Mean proportion recalled for the abstract paragraphs was 54.1% compared to 61.1% for the recall of the concrete paragraphs by the group tested in Experiment 2. Table 4 includes this result along with the recall performance from Experiment 1.

Table 4
Incidental Free Recall Percent Correct under Comprehension Instructions for Monolinguals with Concreteness as a Within versus Between-subjects Factor

<table>
<thead>
<tr>
<th></th>
<th>Abstract</th>
<th>Concrete</th>
<th>Difference</th>
<th>Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Within(^a)</td>
<td>41.9</td>
<td>70.8</td>
<td>28.9</td>
<td>56.4</td>
</tr>
<tr>
<td>Between</td>
<td>54.1</td>
<td>61.1</td>
<td>7.0</td>
<td>57.5</td>
</tr>
</tbody>
</table>

\(^a\) from Experiment 1
In an analysis of variance, this was not a significant difference, $F(1,32) = 2.41, p > .10$. The effect of the order in which the subject received the four paragraphs of each type was not significant, $F < 1$, nor was the interaction of order and concreteness, $F(3,32) = 1.21, p > .10$.

An ANOVA comparing the overall level of recall for the monolingual subjects in Experiment 1 (within-subjects concreteness) to the level of recall for subjects in Experiment 2 (between) showed that the two groups were not significantly different in overall recall, $F(1,78) = 1.73, p > .10$, yet an overall concreteness advantage remained, $F(1,78) = 37.44, p < .001$. There was a significant interaction between concreteness and group from Experiments 1 and 2, $F(1,78) = 12.30, p < .001$. This suggests that despite the concreteness advantage for the subjects in the within-subjects design, both groups show comparable memory overall. The subjects in Experiment 1 appear to be trading off abstract recall for concrete but remaining within the same range of memory performance.

In order to assess the reliability of the effects of concreteness over other materials, proportion free recall was also analyzed, treating paragraphs as the only random effect. In a between-items analysis of variance,
concreteness still did not have a significant impact on performance, $F(1,152) = 2.48$, $p < .12$. The effect of particular paragraph did affect recall, $F(3,152) = 8.51$, $p < .001$, the interaction between paragraph and concreteness was also significant, $F(3,152) = 5.49$, $p < .005$.

**Discussion of Experiment 2**

The main finding of Experiment 2 was the absence of a concreteness advantage when subjects receive only abstract or concrete versions. This supports Hunt & Marschark's (1987) contention that when subjects have only one type of paragraph, they will focus on thematic, shared processing strategies normally used when reading prose material. But it also suggests that the retrieval of the contents of a single paragraph versus the entire set of four at once can alter the effects of concreteness for recall of prose. For prose, the usual advantage of concrete materials is not found. It is likely that the item-specific, distinctive processing necessary to promote better memory for concrete stimuli is not a likely strategy for the reader. Strategy selection can be altered depending on the details of the retrieval procedure.
CHAPTER IV

EXPERIMENT 3

It now seems possible that concreteness effects are tied to differences in strategy use. The inherent "imagery" value of the words within the paragraphs of each type are obviously not changing across Experiments 1 and 2, yet in one case, concrete materials show a memorial advantage and in the other, they do not. The bilinguals in the first experiment did not show a concreteness advantage, even though the monolinguals did in that within-subjects design.

Experiment 3 was designed to induce a concreteness advantage in the bilingual group by promoting attention to the differing concreteness of the key content words in the two types of paragraphs. Since bilinguals did not show an advantage for concrete material, they were asked to rate the imagery value of the sentences within a given paragraph in order to make more salient the difference between the two types. It is important to note that this manipulation should tend to exaggerate item-specific processing over relational, and therefore not only emphasize "imagery" but also distinctive encoding by sentence rather than paragraph. This may serve to dampen
overall levels of recall as well as to promote an advantage for concrete stimuli.

Imagery rating as an orienting task, in comparison to comprehension, may influence the responses subjects subsequently give on self-report estimates of their imagery ability. Rating the imagery value of the sentences in the incidental free recall task may predispose the subject to notice more carefully his or her own imagery ability than when imagery is not mentioned beforehand. Therefore, these measures of individual differences in imagery ability will also be collected as in Experiment 1.

Method

The procedure for Experiment 3 was identical to Experiments 1 and 2, except that this new group of 41 native English bilinguals rated the imagery value (see Paivio, Yuille & Madigan, 1968) of each sentence in each paragraph. Subjects were told to read each sentence within a paragraph, one at a time, and rate the imagery value from 1 (very low) to 7 (very high). The subjects were told to use the entire scale when appropriate, not to worry about using some numbers more often than others, and most importantly, to give their "most honest and consistent" ratings. They were given two example
sentences: low imagery—The truth and beauty of psychology was apparent to everyone; high imagery—The twinkie fell to the floor and the yellow dog devoured it immediately. Subjects had the same amount of time with each paragraph as in the previous experiments and were asked to write their ratings for each sentence on a separate sheet of paper. After all subjects had rated the sentences in the paragraphs (two of each concreteness type) the rating were collected and the subjects were asked as in the preceding experiments to recall as much as possible in any order from the four paragraphs they had just seen. The ratings the subjects gave were collected before the incidental free recall task. As in the previous experiments, subjects were then given the self-rating questionnaires (minus the VVQ, because of time constraints) and the space relations test.

Results

In contrast to the bilinguals in Experiment 1, the concrete paragraphs did show a significant advantage over the abstract for these bilingual subjects when they were asked to rate the imagery value of the sentences within the paragraphs as opposed to reading them for comprehension. Mean incidental free recall performance was 21.7% for the abstract paragraphs and 44.3% for the concrete, $F(1,37) = 52.35$, $p < .001$. As in all previous
analyses, the effect of order was not significant, $F(3,37) < 1$, nor was the interaction, $F(3,37) = 2.01, p > .10$. Table 5 presents free recall performance under comprehension instructions (from Experiment 1) and compares this to recall for bilinguals under imagery rating instructions (Experiment 3).

Table 5
Incidental Free Recall Percent Correct under Comprehension Versus Imagery Rating Instructions for Bilinguals

<table>
<thead>
<tr>
<th></th>
<th>Abstract</th>
<th>Concrete</th>
<th>Difference</th>
<th>Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Comprehension$^a$</td>
<td>51.3</td>
<td>57.9</td>
<td>6.6</td>
<td>54.6</td>
</tr>
<tr>
<td>Imagery rating</td>
<td>21.7</td>
<td>44.3</td>
<td>22.6</td>
<td>33.0</td>
</tr>
</tbody>
</table>

$^a$ from Experiment 1

The intercorrelations for the individual difference measures from the bilinguals after imagery rating instructions reveals only one significant relationship between the VVIQ and the TVIC $r = +.70$. The TVIC and abstract recall performance were significantly correlated, $r = +.36, p < .01$, as were the TVIC and concrete, $r = +.30, p < .06$. The VVIQ and concrete recall were also correlated, $r = +.40, p < .01$. As before, signs of the correlations with the self-report measures are reversed because of the fact that a low score reflects high ability. Again, too, ratings for the self-report measures could be from 1 (high ability) to 5
(low ability), and SRA percent correct was out of a total possible score of 30. Recall scores, ABST and CONC, were out of a possible score of 100 percent. Table 6 presents the correlation matrix.

Compared to the correlation matrix of Experiment 1 for bilinguals (see Table 3) who rated their imagery ability after comprehension instructions versus those bilinguals who rated their ability after imagery rating instructions (see Table 6), more correlations with free recall performance were seen with the latter. Correlations with measures of imagery vividness and controllability correlated significantly with recall performance after imagery rating instructions, but not after comprehension, and with spatial accuracy performance after comprehension, but not imagery rating instructions.
Table 6
Bilinguals' Means, Standard Deviations, Ranges, and Intercorrelations for Self-reported Visual-imagery, Spatial Ability, and Prose Recall Under Imagery Rating Instructions

<table>
<thead>
<tr>
<th></th>
<th>VVIQ</th>
<th>TVIC</th>
<th>SRA</th>
<th>SRT</th>
<th>ABST</th>
<th>CONC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>2.06</td>
<td>2.00</td>
<td>20.1</td>
<td>12.8</td>
<td>21.7</td>
<td>44.3</td>
</tr>
<tr>
<td>Std Dev</td>
<td>.70</td>
<td>.53</td>
<td>5.1</td>
<td>2.1</td>
<td>14.8</td>
<td>18.6</td>
</tr>
<tr>
<td>Range</td>
<td>1-4</td>
<td>1-5</td>
<td>10-29</td>
<td>8-15</td>
<td>0-75</td>
<td>0-75</td>
</tr>
</tbody>
</table>

Intercorrelations

<table>
<thead>
<tr>
<th></th>
<th>VVIQ</th>
<th>---</th>
<th>.70**</th>
<th>.15</th>
<th>.03</th>
<th>.17</th>
<th>.40*</th>
</tr>
</thead>
<tbody>
<tr>
<td>TVIC</td>
<td>---</td>
<td>.20</td>
<td>.07</td>
<td>.36*</td>
<td>.30*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SRA</td>
<td>---</td>
<td>-.16</td>
<td>-.11</td>
<td>.21</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SRT</td>
<td>---</td>
<td></td>
<td>.02</td>
<td>.05</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*p<.05, **p<.0001

A final correlation was calculated in order to determine the relationship between the imagery ratings that served as the orienting task for the subjects' incidental free recall and their performance on abstract versus concrete paragraphs. If subjects rated the paragraphs that they remembered better as more concrete, then this would add evidence to the claim that concreteness was in fact aiding memory even for prose information. There appears to be no reliable correlation between each subject's absolute value difference between concrete versus abstract imagery ratings and their
performance difference between the two types of paragraphs, $r = +.27$, $p < .10$.

**Discussion of Experiment 3**

Clearly, asking bilingual subjects to rate the imagery value of sentences that make up the concrete and abstract paragraphs draws their attention to both concreteness and more item-specific processing than do comprehension instructions. Rating imagery value produces a poorer incidental recall performance overall and a 2:1 advantage for concrete stimuli. This overall poorer performance allows rejection of the argument that the concreteness advantage for monolinguals in Experiment 1 had anything to do with higher levels of recall, since the present results are similar in range to those found in Marschark's results. Also, the increased incidence of correlations between recall and imagery measures argues against considering the task as one of simply focussing the subjects on isolated sentences. Subjects are now in some way linking what they do in rating their imagery ability with how well they perform on recall of abstract versus concrete paragraphs, and this link is more apparent when the preceding orienting task is rating sentence imagery value than when it is merely reading for comprehension.
CHAPTER V
EXPERIMENT 4

Method and Results

The procedure followed in Experiment 4 was identical to the one in Experiment 3, except that a new group of 43 monolingual subjects (29 women and 14 men) from an Introduction to Cognitive Psychology course participated in partial fulfillment of a course requirement. As in Experiment 3, subjects rated the imagery value of the sentences in the four paragraphs, performed in an incidental free recall for the paragraphs, and then completed two self-ratings of imagery ability (VVIQ, TVIC) and the space relations test (SRA, SRT).

As with the bilingual group in Experiment 3, monolingual subjects showed a 2:1 advantage for concrete paragraphs relative to abstract (37.5% and 18.2%, respectively). In an analysis of variance with concreteness as a within-subjects factor, this proved to be a significant difference, $F(1, 85) = 27.50$, $p < .0001$. As in all previous analyses, order of paragraph presentation was not a significant effect, $F(3, 85) < 1.0$, nor was the interaction between order and concreteness, $F(3, 85) = 1.49$, $p > .10$. The overall mean performance
for this group was 27.8%, compared to 33.0% for the bilinguals in Experiment 3. In an analysis of variance treating concreteness as a within-subjects factor and language as a between-subjects, this was not a significant difference, $F(1,167) = 3.02, p < .10$. Since recall performance was at least as high for the bilinguals as for the monolinguals, the possibility of the subjects in Experiment 4 (who were more advanced students) being more "skillful" in general than the sample taken from an introductory level group seems unlikely.

Concreteness was a significant factor for both the bilinguals of Experiment 3 and the monolinguals of Experiment 4, $F(1,167) = 70.85, p < .0001$ and the interaction between language (experiment) and concreteness was not significant, $F(1,167) = .45, p < .51$. Table 7 contains the incidental free recall percent correct under imagery rating instructions for the monolingual group (Experiment 4) compared to the bilinguals of Experiment 3.
Table 7
Incidental Free Recall Percent Correct under Imagery Rating Instructions for Monolinguals versus Bilinguals

<table>
<thead>
<tr>
<th></th>
<th>Abstract</th>
<th>Concrete</th>
<th>Difference</th>
<th>Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Monolinguals</td>
<td>18.2</td>
<td>37.5</td>
<td>19.3</td>
<td>27.9</td>
</tr>
<tr>
<td>Bilinguals^a</td>
<td>21.7</td>
<td>44.3</td>
<td>22.6</td>
<td>33.0</td>
</tr>
</tbody>
</table>

^a from Experiment 3

The intercorrelations for the individual difference measures from the monolinguals after imagery rating instruction reveals only one significant relationship, between the VVIQ and the TVIC \( r = +.30 \). This can be compared to the same results for the bilinguals in Experiment 3, where this correlation was +.70, and again the only significant one. To determine whether these two correlation coefficients are significantly different, a Fisher's transformation to \( z \) was performed. The difference between \( z \)'s (.31 and .87, respectively) deviates from a difference of 0.0 to the extent of 2.48, which means that it is significant at the .05 level.

In contrast to the bilingual group, the only significant correlation between any individual difference measure and recall performance was SRT and CONC, \( r = - .37 \). Table 8 contains the correlation matrix for Experiment 4.
Table 8
Monolinguals’ Means, Standard Deviations, Ranges, and Intercorrelations for Self-reported Visual-imagery, Spatial Ability, and Prose Recall Under Imagery Rating Instructions

<table>
<thead>
<tr>
<th></th>
<th>VVIQ</th>
<th>TVIC</th>
<th>SRA</th>
<th>SRT</th>
<th>ABST</th>
<th>CONC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>2.06</td>
<td>1.87</td>
<td>17.9</td>
<td>12.5</td>
<td>18.2</td>
<td>37.5</td>
</tr>
<tr>
<td>Std Dev</td>
<td>.58</td>
<td>.44</td>
<td>4.65</td>
<td>1.52</td>
<td>20.3</td>
<td>25.0</td>
</tr>
<tr>
<td>Range</td>
<td>1-5</td>
<td>1-3</td>
<td>11-29</td>
<td>9-14</td>
<td>0-90</td>
<td>0-90</td>
</tr>
</tbody>
</table>

Intercorrelations

<table>
<thead>
<tr>
<th></th>
<th>VVIQ</th>
<th>TVIC</th>
<th>SRA</th>
<th>SRT</th>
<th>ABST</th>
<th>CONC</th>
</tr>
</thead>
<tbody>
<tr>
<td>VVIQ</td>
<td>---</td>
<td>.30*</td>
<td>.16</td>
<td>.04</td>
<td>.05</td>
<td>.03</td>
</tr>
<tr>
<td>TVIC</td>
<td>---</td>
<td>.13</td>
<td>.04</td>
<td>.14</td>
<td>.04</td>
<td></td>
</tr>
<tr>
<td>SRA</td>
<td>---</td>
<td></td>
<td>-.10</td>
<td>.09</td>
<td>.02</td>
<td></td>
</tr>
<tr>
<td>SRT</td>
<td>---</td>
<td></td>
<td></td>
<td>.07</td>
<td></td>
<td>-.37*</td>
</tr>
</tbody>
</table>

*p<.05.

Discussion of Experiment 4

The free recall performances of the monolinguals mirror those of the bilinguals. A 2:1 concreteness advantage for both language groups is found when subjects rate the imagery value of the sentences in the paragraphs as an orienting task for the subsequent incidental memory task. Under conditions of comprehension orienting instructions, monolinguals show a concreteness advantage whereas bilinguals do not. But under the conditions of imagery rating instructions, the two language groups perform in an equivalent manner. It appears that under imagery rating instructions for both language groups,
subjects are treating the sentences as if they were in random order. This would compare favorably then with Marschark's (1985) finding of a concreteness advantage when sentences are not processed relationally. There is still no existing rationale for the language group dissociation when subjects process the paragraphs by relating the sentences together to form a coherent context. In the general discussion, this question will be addressed in more detail.

It is more difficult to find similarities across Experiment 3 and 4 for the individual difference measures. The only significant intercorrelation among these measures for both language groups (i.e., both Experiments 3 and 4) was between the VVIQ and the TVIC, but with a total sample size of 84, the correlation of $r = +.70$ for the bilinguals is significantly higher than the $r = +.30$ for the monolinguals. The bilinguals are linking their self-rated vividness and controllability of imagery more closely than the monolingual group.

A particularly apparent difference between individual difference correlations in Experiments 3 and 4 lies in the lack of such correlations with recall performance for the latter experiment. In Experiment 3, both the VVIQ and the TVIC correlate significantly with concrete recall, while the TVIC also correlates with
abstract recall. In Experiment 4, the only reliable correlation is the one between SRT and concrete recall. That is, the longer an individual took to complete the space relations test of 30 items, the less well they had remembered the concrete paragraphs. This is the only significant relationship with the variable SRT.
CHAPTER VI
GENERAL DISCUSSION

Five main findings emerge from the above results. First, when concreteness is treated as a within-subjects factor, this manipulation does produce a significant advantage for monolingual subjects, even at the level of connected prose material. Second, when using the identical procedure with bilinguals, these subjects do not show a concreteness advantage. Third, for both language groups, intercorrelations among visual-spatial imagery and between these self-report measures and free recall performance are not all that strong and appear to be similar for both groups of subjects when the preceding task is comprehension of the paragraphs. Fourth, when treating concreteness as a between-subjects factor, monolinguals, like bilingual subjects, no longer show a concreteness advantage. And finally, when the "concreteness" of the words within the sentences is made more salient by asking subjects to rate sentences on this dimension, a concreteness advantage occurs for both language groups. Also, significant correlations between self-ratings and free recall are much more apparent in
the imagery rating condition for the bilinguals than for the monolinguals.

These five findings point to a focus on strategy differences that are not so much related to individual self-reports of visual imagery ability or spatial aptitude, but to such factors as whether a subject is a speaker of one or two languages (Experiments 1, 3 and 4), whether concreteness is manipulated within or between subjects (Experiment 2), and specific orienting tasks (Experiments 3 and 4). The following sections include more detailed discussion of the above five findings and conclude by describing some possible interpretations and extensions.

**Manipulating Concreteness Between or Within Subjects**

Similar to previous tests of the concreteness effect in memory for word lists, a significant difference was found between free recall performance for concrete versus abstract paragraphs using a within-subjects design, with the former having an almost 2:1 advantage over the latter. However, when concreteness was a between-subjects factor, the concreteness effect disappeared. Marschark (1985) reported an equivalent level of performance for both types of paragraphs, but did not initially qualify the results to include only conditions where subjects did not receive paragraphs of
mixed concreteness. Obviously, the manipulation of a within- versus between-subjects factor of concreteness plays an important role in the ability of subjects to differentially remember concrete and abstract prose material. Since the imagery values of the words within the paragraphs do not change as a function of the within- versus between-subjects manipulation, these results are probably better explained in terms of processing rather than representation [see Marschark, Richman, Yuille, & Hunt (1987)].

There has been evidence of an interaction between within- versus between-subjects design and the distinctiveness of the stimuli in a study by Hunt & Elliot (1980). Hunt & Elliot found that irregular orthographic patterns rated as distinctive facilitated memory only when presented in the context of orthographically regular words. Distinctive features of words are those shared by few other words. For example, an orthographically irregular word would be one such as "phlegm" and an orthographically regular word would be one such as "primate". Hunt & Elliot's result that distinctive processing will have its greatest impact in a within-subjects design could serve as partial explanation for the findings of Experiments 1 and 2. Concrete words could be said to be more distinctive because they have
typically been found to be more specific to one context. Rating the distinctiveness of a word sounds procedurally similar to the context-availability rating that was asked of subjects in the Schwanenflugal & Shoben study mentioned earlier.

Recall that in Experiment 1 (mixed concreteness design) monolingual subjects remembered the concrete paragraphs better than the abstract, while in Experiment 2 (single concreteness design) memory was equivalent. Hunt & Marschark (1987) have suggested that subjects attend to relational or shared information during processing of paragraphs of one type of concreteness. Since they are processing prose material, relational strategies are appropriate and obvious in light of comprehension as a goal and subjects may not even be able to notice concreteness. However, when subjects receive paragraphs of both types of concreteness, the distinctive information inherent in concrete words is made more salient by proximity to abstract words. If a concrete paragraph is in direct comparison with words that are less so, then the concrete paragraph will be more likely to receive unit-specific, distinctive information as well as relational than will the abstract. Thus concrete paragraphs presented with abstract may gain the benefits of both types of processing providing them with more than
one retrieval route and subsequently better recall in a mixed design.

**A Question of Retrieval Versus Encoding Differences**

Total retrieval attempts of information from all abstract and concrete paragraphs presented rather than a read paragraph—recall paragraph sequence also seems to influence the concreteness advantage. Besides the present study, Bunn varied concreteness within-subjects and then considered the data from only the first presentation for all subjects. In this approximation of a between-subjects manipulation, no concreteness effect was obtained in either gist or verbatim scoring of idea units recalled. When looking at the complementary, second-presented material set, however, a concreteness effect was obtained. In contrast to both of these results, the present study revealed a concreteness advantage when controlling for concreteness of the first presentation and practice paragraph. However, this result must be tempered by the fact that subjects in the present study made a recall attempt only after having studied all paragraphs, while in the Bunn study, subjects recalled a paragraph directly after each was presented. The above finding points to the idea that retrieval processes rather than those at encoding may be more important in determining whether a concreteness effect is
obtained. This is a critical distinction to make since in previous interpretations of the concreteness effect (for example, Marschark, 1985) the focus was on possible encoding strategy differences induced by high- and low-imagery materials. Further studies must dissociate the effects of manipulating concreteness within- versus between-subjects from the effects of the retrieval procedure. Bunn's result, along with the analysis of recall adjusting for zero recalls in Experiment 1 (see Results), suggest that perhaps concrete concepts, given similar retrieval success, are only slightly more retrievable than abstract ones.

Effects of Bilingualism on Prose Recall

The main difference between monolinguals and bilinguals under comprehension instructions (Experiment 1) was the lack of a concreteness advantage in an analysis by items for the bilingual group. This is in contrast to previous research (Ransdell & Fischler, 1987; Miljkovitch 1980) which seemed to show that compared to monolinguals, bilingual speakers were better able to take advantage of imagery inherent in stimuli. Perhaps this result supports Marschark's contention that relational strategies play a more important role than do imaginal strategies at this level of analysis. Alternatively, that bilingual subjects, like subjects confronted with
only one type of concreteness, may not show the concreteness effect because they simply do not notice concreteness as much when they normally confront stimuli in their everyday reading experiences. Perhaps the bilingual must rely on relational processing more than on item-specific, even when viewing concrete and abstract material together. This follows from the results of Experiment 3, where bilingual subjects showed a distinct advantage for concrete paragraphs when their attention was drawn to this dimension.

Recall that bilingual subjects were not aware that they were participating in the experiment because they were speakers of more than one language. This implies that overt second language activation was kept to a minimum or at least of some latent nature. Future studies will need to explicitly look at concreteness effects under conditions of high and low second language activation to discover the reason behind language group differences in recall.

Visual Imagery Differences: Monolinguals

Further partial support for the idea that visual imagery per se does not facilitate memory for concrete versus abstract prose can be inferred from lack of strong correlations between imagery measures and recall performance. Intercorrelations among the imagery
measures were in agreement with past findings and relationships between a subjective verbal report of spatial control (TVIC) and the VVIQ and VVQ were found. It is interesting that the only significant correlations between recall and imagery measures after comprehension instructions were between the only performance measure, accuracy on space relations and abstract and concrete recall.

This pattern of intercorrelations changes when the orienting task before the incidental memory task is imagery rating of sentences within the paragraphs rather than comprehension. The self-ratings are still interrelated, but recall correlates with SRT, not with SRA, as in the comprehension condition of the first experiment. Possibly the most important, and surprising result from the correlation patterns under the two orienting tasks, is how readily relationships between recall, self-ratings of imagery ability and spatial aptitude change, given the type of task the subject performs preceding incidental recall.

**Visual imagery differences: Bilinguals**

The correlations among the bilingual subjects were suggested to have more influence in predicting recall performance than for the monolinguals, and it was also hypothesized that bilinguals would report higher visual
imagery scores. Under imagery rating conditions, subsequent self-ratings of imagery ability were found to correlate with recall in more instances for the bilinguals than for the monolinguals. Under comprehension conditions, the pattern of correlations were not significantly different between the two language groups. In this case, scores for the individual difference measures were nearly identical for both language groups. As to the second hypothesis, there is no evidence that bilinguals rate themselves any differently on imagery measures than monolinguals.

Bilinguals' self-reported imagery ratings do not correlate with their prose recall performance when the preceding task is to comprehend abstract and concrete paragraphs. But these variables do correlate after imagery rating instructions. This can be taken as evidence that concreteness does not play as large a role in comprehension of prose as it does in rating imagery value, as is reflected in incidental memory performance. Intercorrelations are nearly identical for the monolingual subjects, while their recall performance was not the same, namely, a concreteness advantage under within-subjects manipulation of concreteness for the monolingual group. This adds evidence that imagery per se is not at the heart of the concreteness effect at the
level of connected discourse. That is, visual imagery ability for the two language groups does not differ, while recall performance and evidence of a concreteness advantage are different for speakers of one and two languages.

This raises interesting possibilities for the different performance strategies between monolingual and bilingual speakers. Given that the two language groups rate themselves similarly on visual-spatial imagery dimensions, a dissociation between imagery strategy use and ability may be driven by the bilingual's capabilities in accessing two linguistic codes. Perhaps at the level of relational processing inherently present in prose stimuli, the salience of the concreteness of the words within the paragraphs is even less important to bilinguals than it is to monolinguals. Recall that the bilingual group performed more like the monolingual subjects who received paragraphs of only one type. Bilinguals, and monolinguals in a between-subjects manipulation of concreteness, did not remember concrete paragraphs any better than abstract. Since the rationale for the within- versus between-subjects concreteness difference implied that concreteness was more noticeable in the context of words that were also quite abstract, then bilinguals might simply be less attentive to
concreteness. This result, along with the individual difference correlations with recall, points to the contention that "concreteness" is not the same construct as "imagery" at the level of prose processing and that possibly neither of these stimulus characteristics underly strategies in remembering prose. Strategies based on the relating of the ideas within a coherent contextual theme may override strategies based on the concreteness of stimuli or on individual propensities to use visual-spatial imagery.

**Future Research and Extensions**

Several main questions are raised by the results of the present set of experiments. The determination of whether concreteness effects are more of a retrieval or encoding phenomenon is a central issue. Past research has been equivocal on this question and has not directly tested it. There is evidence from the present experiment that controlling for whether any of a paragraph is retrieved, memory for the two types is not as discrepant as when instances of "zero" recall are included. The greatest incidence of "zero" recall (that is, when none of the paragraph is retrieved) was from the abstract paragraphs in the within-subjects experiment. Partially cueing subjects by prompting them with the first few words of a given abstract or concrete paragraph could
give a clear picture of the nature of concreteness effects unconfounded by the relative retrieval difficulties of abstract paragraphs.

One way to determine the effects of context specificity and multiple retrieval attempts on the concreteness effect would be to systematically vary whether a given concrete or abstract content word is initially encoded in the context of a word list or within a paragraph. Subjects would receive a word in a list on the first trial, followed by a memory test, and then would receive that same word in the second trial embedded within the context of a paragraph, again followed by a recall test. Some subjects would receive the words in paragraph context first, and then in list context. The effects of the repetition of the word and its subsequent context by order recall would be compared across subjects receiving mixed versus unmixed concreteness of materials, and across subjects who retrieved after only the first trial (list or paragraph), the second trial, or after both. Context availability (Schwanenflugel & Shoben, 1983) could then be explicitly specified, as well as number of retrieval attempts. Subjects in each condition could subsequently be asked to rate the concreteness and context specificity of all stimulus words, before, after, or at both of these times relative to the recall trials.
A further study of the "suggestability" of individual difference measures of imagery is important since the orienting task can so easily alter self-reports and spatial performance. Many studies whose goal is to correlate imagery ability and recall performance do not consider the pre-rating environment of the individual difference testing situation. Further studies must explicitly outline the entire experimental episode, including what the subjects do before rating their abilities or performing some spatial task. This holds even when subjects are unaware of the connection between recall and ability tests. Subjects could be given individual difference tests in advance of any recall performance. Alternatively, the orienting task could be systematically altered to determine the pattern of incidental effects of recall attempts of concrete and abstract material on imagery ability measures.
APPENDIX A

FOUR PARAGRAPH PAIRS USED AS STIMULI

Once while working in a guided missile plant, Andy Gardner was approached by a bearded foreign agent. The agent, who was known for his disguises as a young hippie, offered him a job building atomic bombs. Andy was offered a new sportscar and a huge mansion. Also, a beautiful blonde assistant and a modern laboratory were at his disposal. Although tempting, his nagging wife would not let him take the job. The whole affair was reported to his overweight supervisor, who just thoughtfully nodded his head.

Once while working on a secret defense project, Andy Gardner was contacted by a certain foreign government. The government, which was known for its activities in international espionage offered him a job designing weapons systems. Andy was offered a high salary and short hours. Also, a competent research staff and unlimited funds were at his disposal. Although tempting, his patriotic conscience would not let him take the job. The whole affair was reported to the security division, which simply appeared to ignore the incident.
As head of the marines, Ted Felton made an embarrassing mistake concerning some black soldiers stationed in a hot and steamy town in Texas. A riot broke out when the soldiers were accused of beating three old ladies. Although a colonel present at the scene disputed the charge, Felton ordered all of the men involved thrown in jail. Later, the elderly women testified that they had not actually seen their attackers. Felton admitted his mistake and personally opened the door to the jail.

As head of the military, Ted Felton made an unwise decision concerning some minority group servicemen stationed in a quiet and remote part of Texas. A disturbance occurred when the servicemen were accused of assaulting several local residents. Although the evidence available at the time disputed the charge, Felton ordered all those involved put in custody. Later, the involved citizens realized that they had not actually seen their attackers. Felton realized his mistake and immediately ordered the men to be freed.
Early in his college career, John Williams was chosen by his football coach to receive a trophy from a professional team. The ceremony proved to be crowded and was prolonged over a two hour period. John was excited however to receive awards for scoring and sportsmanship from their star player. Further, his large size and fast speed helped to improve opinions of his college's players. Williams went on to become a star center and gained an international reputation.

Early in his college career, John Williams was chosen by his department to receive a scholarship to a foreign university. The program proved to be difficult and was prolonged over a several year period. John was successful however in obtaining degrees in philosophy and humanities from the institution. Further, his keen interest and tactful personality helped to improve opinions of Canadian students. Williams went on to become a brilliant philosopher and gained an international reputation.
Greg Jackson had a large file-folder on each new student in his class. With this information he was able to deal with discipline problems in a way that won the admiration of younger teachers. Shortly after Greg began teaching in the woodworking shop, a youth threw a glass bottle at his new station wagon, which was painted fire-engine red. The student accused Jackson of giving an unfair exam. During the next class, however, no other students complained when asked. Later, in fact, the skinny student was beaten-up by the other pupils.

Greg Jackson had a thorough knowledge of each new development in his field. With this information he was always able to interpret relevant findings in a way that won the admiration of other researchers. Shortly after Greg became interested in the history of science, a colleague wrote a bitter critique concerning his earlier theoretical work, which was concerned with language development. The essay accused Jackson of ignoring some important facts. Over the next month, however, no other critics responded when queried. Later, in fact, the offensive essay was ignored by other writers.
APPENDIX B
AN EXAMPLE FROM EACH INDIVIDUAL DIFFERENCE MEASURE

VVIQ (Vividness of Visual Imagery Questionnaire)
Think of some relative or friend whom you frequently see (but is not with you at present) and consider carefully the picture that comes before your mind's eye.
1. The exact contour of face, head, shoulders and body.
2. Characteristic poses of head, attitudes of body, etc.

TVIC (Test of Visual Imagery Control)
1. Can you see a car standing in the road?
2. Can you see it in color?
3. Can you now see the car lying upside down?

VVO (Verbalizer-Visualizer Questionnaire)
1. I enjoy doing work that requires the use of words.
2. I don't believe that anyone can think in terms of mental pictures.

SR Differential Ability Test (Space Relations)
For each of the following "2-D" patterns on the left, decide which of the "3-D" patterns on the left can be made from it.
APPENDIX C

LANGUAGE HISTORY QUESTIONNAIRE

age    sex
years in U.S.    years in U.S. schools

1) What language would you consider as your native language?
   How many years have you spoken this language?
   Do one or both parents or caretakers speak this language?
   Was this language learned as your first language when
   you were first learning to speak?

2) What language would you consider as your second language?
   How many years have you spoken this language?
   Was this language learned in the home or in school?

Please rate your language ability according to the following scale.
1-very fluent  2-fluent  3-weakly fluent
4-barely fluent  5-not fluent

written   spoken

Native language
Second language
REFERENCES


BIOGRAPHICAL SKETCH

Sarah Ellen Ransdell has spent the last 20 years, out of her 26, as a student. And like all good things, this is about to come to an end, at least formally. The author attended University of Kentucky from 1979 to 1983 and received her Bachelor of Arts degree as a psychology major. She then spent the next four years as a graduate student in cognitive psychology at the University of Florida where she, at the time of this writing, wildly anticipates a doctoral degree in August of 1987.

Sarah will become Assistant Professor of Psychology at the University of Maine in Orono in the fall of 1987. In this, her first "real" job, she will focus on starting a computer lab for psychology undergraduates. She will be teaching "Maine"iacs the principles of psychology and will be doing some research. Perhaps most importantly, she will be establishing the career that has been her goal for most of the last decade.
I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Irâ S. Fischler, Chairperson
Professor of Psychology

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

C. Michael Levy
Professor of Psychology

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Richard A. Friggs
Professor of Psychology

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Walter Cunningham
Professor of Psychology
I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

William J. Sullivan  
Associate Professor  
of Slavic and Germanic Languages and Linguistics

This dissertation was submitted to the Graduate Faculty of the Department of Psychology in the College of Liberal Arts and Sciences and to the Graduate School and was accepted as partial fulfillment of the requirements for the degree of Doctor of Philosophy.

August, 1987

Dean, Graduate School