Verbal mediation of visual memory on the Continuous Visual Memory Test

Sean P. Whalen
The University of Montana
Permission is granted by the author to reproduce this material in its entirety, provided that this material is used for scholarly purposes and is properly cited in published works and reports.

** Please check "Yes" or "No" and provide signature **

Yes, I grant permission
No, I do not grant permission

Author's Signature ______________________________

Date 5/10/99

Any copying for commercial purposes or financial gain may be undertaken only with the author's explicit consent.
Verbal Mediation of Visual Memory
on the Continuous Visual Memory Test

By
Sean P. Whalen
B. S., Psychology, Pacific Lutheran University

Presented as partial fulfillment of the requirements for the
Degree of Master of Arts
University of Montana
1999

Approved by

Committee Chair
Dean of the Graduate School

Date
Nonverbal memory testing is an integral component of clinical neuropsychological evaluation. The idea of nonverbal memory is supported both by research on hemispheric asymmetry and by some factor analyses of memory test batteries. The construct of nonverbal memory has received inconsistent support from empirical research. One significant, potential confound of nonverbal memory testing is verbal mediation or contamination of visual stimuli. Investigation of several nonverbal memory tests indicates that test items are susceptible to verbal contamination.

The Continuous Visual Memory Test (CVMT, Trahan & Larrabee, 1988) is a relatively new and little researched test of nonverbal memory. The CVMT utilizes a recognition format and uses complex geometric figures as visual stimuli, and stimulus duration is brief. These characteristics are proposed as means to minimize confounds, including verbal contamination of the test. The verbal characteristics of the test items, however, have not been empirically examined.

In the present study two experiments were performed. In the first, the verbal characteristics of the CVMT were investigated. 51 students from an introductory psychology class were administered the CVMT and asked to assign verbal labels to the items. Subject responses were recorded and several indices of verbal loading of items were calculated using those responses. Results indicate that the CVMT items are susceptible to verbal encoding, and that some classes of visual stimuli are more prone to verbal labeling than others.

In the second experiment the relationship between verbal load of test items and subject performance was examined. 51 volunteer students from an introductory psychology class participated. Each was administered the CVMT according to standard instructions. Verbal load of the stimulus items was not consistently related to subjects' ability to correctly identify target items at delayed recognition. However, verbal load of item classes did relate to subject performance using one index of verbal contamination. In addition, the overall performance of subjects in experiment 1 and experiment 2 was compared. Results indicate that there were no significant differences in performance at delayed recognition between the two groups. Implications for interpretation of CVMT performance, future research, and test development were discussed.
Table of Contents

Abstract ii
List of Tables and Figures iv
Introduction 1
Experiment 1
  Method 14
  Results 19
Experiment 2
  Method 26
  Results 27
Discussion 33
Appendix A
  Neuromedical Screening Form 41
Appendix B
  Informed Consent Form 42
Appendix C
  Debriefing Script 43
Appendix D
  Instructions Script, Experiment 1 44
Appendix E
  Instructions Script, Experiment 2 45
References 46
List of Tables and Figures

Tables

1. Association, Content, and Heterogeneity of Response (H-item) for Individual Items  20
2. Mean A Values and Grouping for Each of the 7 Item Classes  21
3. Mean C Values and Grouping for Each of the 7 Item Classes  22
4. Mean H-item Values for Each of the 7 Item Classes  23
5. H-class Values and Rankings for Each of the 7 Item Classes  24
6. H-class Values and Rankings for Each of the 7 Item Classes  25
7. Average Performance at Delay for Each Item Class  28

Figures

1. Examples of Stimulus Items  18
2. Subject Performance at Delay by Verbal Load of Target Items  29
3. Subject Performance at Delay by Verbal Load of Item Class  32
Introduction

Verbal and Nonverbal Memory

Within clinical neuropsychology, the measurement of memory is viewed as a common and fundamental aspect of assessment. In fact, Lezak (1995) describes memory and learning as forming one of the four basic classes of cognitive functioning. Furthermore, most clinicians make a distinction between verbally mediated- and nonverbally mediated memory. Making such distinctions is helpful both in assessing lesion location as well as describing the functional aspects of cognition.

The construct of nonverbal memory is not a new one. Spearman (1932) identified a number of factors that related to memory, including sensory memory, verbal-, and nonverbal memory. Memory for words or sentences was included in the verbal memory factor. Nonverbal memory included memory for digits, non-linguistic symbols, and mathematical operations. Nonverbal memory was further divided into visual and auditory modalities. These different modes of processing contributed to the sensory memory factor, a short lived, multi-modal memory system.

Guilford (1967) described memory as one of several domains of intelligence. Memory was further divided into several classes, referred to as aspects of retentivity. Factor analysis of data gathered by previous memory researchers as well as Guilford's studies with the Army Air Forces supported the constructs of figural, symbolic, and semantic memory (Guilford & Lacey, 1947; Christal, 1958; Tenopyr, 1966). Figural memory tests included map memory in which subjects were shown map sections and asked to remember landmarks and spatial relations. Tests of form and color memory also loaded on the figural memory factor. Semantic memory tests included memory for names, memory for instructions, and memory for limericks. All of these are variations of tests that require the subject to retain and repeat verbal material presented visually and/or auditorially. Symbolic memory was weakly supported, including tests of letter or digit span and recall of lists of nonsense words or syllables.
Hemispheric Differences in Verbal and Nonverbal Cognition

The idea of separating verbal and nonverbal cognition, including memory, is supported by lateralization of function within the cerebral hemispheres (Hellige, 1990; Bryden, 1983; Squire & Slater 1978; Tokar, Matheson, & Haude 1989; Gevins, Cutillo, & Smith 1995). Hellige (1990) notes that in humans the left hemisphere is strongly dominant in the production of speech, to the degree that an estimated 95% of right-hand dominant individuals have speech production limited to the left hemisphere. The superiority of the left hemisphere in recognition of verbally processed stimuli exists, but with greater variability, and is described as being more a matter of degree than the absolute of language production (Hellige, 1990). The right hemisphere, he notes, does not appear to have as dramatic or complete a dominance over any function as the left does with regard to language. However, a number of studies have produced converging evidence for right hemisphere superiority over the left in visuospatial and manipulospatial processing (for review see Bryden, 1983). Heilman and Valenstein (1993) report, additionally, that unilateral neglect is far more frequent after right hemisphere damage than after damage to the left hemisphere, supporting the notion of right hemisphere dominance in visual and spatial processing.

Evidence for laterality exists in non-human species as well. Hopkins and Morris (1989) demonstrated a left visual field-right hemisphere dominance for visual spatial processing in two language-trained chimpanzees (Pan troglodytes). The chimpanzees used a joystick to manipulate icons on a computer screen for food rewards. Results suggested that both subjects showed a left visual field-right hemisphere bias in performance, especially utilizing reaction time data, suggesting lateralized function for visual-spatial processing in language-trained nonhumans. Additional support for lateral asymmetry in nonhumans is found in macaques. Heffner and Heffner (1984) reported that Japanese macaques are better able to discriminate species specific calls when presented to the left hemisphere. While not actual language, these
calls may be viewed as an analog to human language. These studies demonstrate lateralization effects in nonhumans similar to that of humans discussed above.

**Clinical Assessment of Nonverbal memory**

Clinical memory testing is an important part of a full neuropsychological battery (Lezak, 1995). Laterality information suggests that testing both verbal and nonverbal memory may aid in lesion localization. Clinical testing using both verbal and nonverbal measures may allow identification of laterality of lesions in the brain if one function is disturbed but not the other. Complimentary to this, one can make a case for diffuse damage if both sets of abilities suffer after damage to the brain. Furthermore, understanding the functioning in these two distinct domains helps to form a holistic view of the patients’ level of cognitive functioning, including specific areas of strength and weakness.

In a review of clinical memory testing, Erickson and Scott (1977) report that Wells and Martin published the earliest memory battery in 1923. The test included both visual and verbal modalities and included measures of both remote, previously learned material and short-term, new learning. These authors, however, did not include measures of what we would now call long-term memory of new material. A very high correlation (.81) was found between memory performance and general intelligence as measured by the Stanford-Binet intelligence scale. This should hardly be surprising, however, as the various revisions of the Binet intelligence tests have always included measures of short-term verbal and nonverbal memory (Cohen, Swerdlick, & Smith 1992). In fact, short-term memory is one of the major factors of the test battery, including primarily nonverbal tests such as the Bead Memory and Memory for objects subtests.

One test commonly used to test visual/nonverbal memory is the Rey-Osterreith Complex Figure Test or one of its many variants (Rey, 1941; Osterreith, 1944; Lezak, 1995). Common administration involves having patients copy the figure, then draw the figure immediately from memory as well as at a number of delay intervals. Multiple scoring procedures are available, with appropriate norms. These procedures allow for awarding points for locating details of the figure in correct locations, yielding total recall scores. Loring, Lee, and Thompson (1989) were
unable to differentiate left and right temporal lobectomy patients using total recall scores using
the Rey-Osterreith and Taylor complex figures, although qualitative analysis of errors made did
distinguish the two groups. Thus, while the actual number of errors made did not differentiate
groups, group membership could be identified by the pattern or type of errors made. Loring,
Lee, and Thompson (1989) recommend caution in interpreting results utilizing recall of the
Complex figure. Lezak (1995) notes that poor performance on the Complex Figure may be due
to a number of factors other than nonverbal memory dysfunction, such as: attention deficits,
poor organizational and problem solving strategies, and visual-perceptual deficits. Frontal lobe
patients are noted to perseverate and confabulate, resulting in poor performance. Thus, while
the Rey-Osterreith may be sensitive to brain injury, it is not specific to right hemisphere
dysfunction or nonverbal memory deficits.

The Benton Visual Retention Test (BVRT) (Benton, 1974; Sivan, 1992) is another
common test used to assess nonverbal memory function. Patients view figures and then on
different forms, copy or recall the figures after delay periods. Lezak notes that decreased
performance on the BVRT is not specific to patients with right hemisphere damage, but that
aphasic patients with left hemisphere damage show poor performance, and that scores increase
as language performance improves. Like the Complex Figure, the BVRT tests a variety of
skills, including, but not limited to, visual memory, and has been shown to be a general indicator
of cerebral dysfunction (Lezak, 1995).

Bomstein (1982) tested individuals with clearly documented unilateral brain lesions
using the Wechsler Memory Scale (WMS, Wechsler 1945). Lesions resulted from surgery,
vascular disorders, focal atrophy, or other damage to the brain. Although neurologically
impaired, these individuals did not differ significantly in overall memory performance as
estimated by the WMS memory quotient (MQ). However, patients with right hemisphere
lesions performed significantly worse than patients with damage to the left hemisphere on the
Visual Reproductions subtest, which is described as a test of nonverbal, visual memory. In
contrast, patients with left hemisphere lesions performed significantly worse than right
hemisphere patients on Logical Memory and Paired Associates, both of which are verbal memory tasks.

Russell (1975) revised the WMS to include delayed recall measures for both the Logical Memory and Visual Reproductions subtests. Russell's revised administration and scoring procedures yielded six reliable memory scales: verbal short-term, verbal long-term, verbal % retained, figural short-term, figural long-term, and figural % retained. These measures were sensitive to brain damage and separated those with brain lesions from both normal controls and medical patient controls. Furthermore, individuals with localized right hemisphere lesions performed worse than those with left hemisphere lesions on figural memory scales, a primarily nonverbal memory task. Subjects with specific left hemisphere damage performed worse than right hemisphere patients on verbal measures. Thus the Russell revision of the WMS was demonstrated to be sensitive and specific to brain damage, as well as lateralized lesions.

The Wechsler Memory Scale-Revised (WMS-R, Wechsler, 1987), the most commonly used test battery for the assessment of memory (Robinette, Sherer, & Adams, 1993), includes a number of tests which purport to test nonverbal memory. In the final summary of scores the WMS-R differentiates between verbal memory and visual memory, with index scores for each. These scores sum to form the index for general memory. The Wechsler Memory Scale was formally revised in 1987 (Wechsler, 1987), adding a number of tests and incorporating the delay procedures described and demonstrated by Russell (1975). Visual paired associates were added with both an immediate and delayed recall condition, and learning to 100% performance or 6 trials whichever comes first. These scores were then weighted and added to the Visual Reproductions subtest scores to derive a Visual Memory Index. Clinical findings and factor analysis are equivocal in their support for the added indices (Lezak, 1995). For example, clinicians agree that the multiple index approach is superior to the former MQ derived by the WMS. However, clinical and research findings suggest that the verbal and nonverbal indices are not always sensitive to brain lesions and that the factors themselves are not, upon examination, particularly robust. A study of the WMS-R performed by Chelune and Bornstein (1988) found
that patients with left and right unilateral lesions performed equally well on the visual paired
associates with a very slight and nonsignificant advantage to right hemisphere patients.
Furthermore, factor analysis failed to resolve a verbal-nonverbal factor split. Significant factors
derived were Attention and Concentration, Immediate Memory, and Delayed Recall. Roth, et
al (1990) performed a factor analysis of the WMS-R and also failed to resolve a visual-verbal
factor discrimination that improved fit over the factors derived by Chelune and Bornstein.
However, Robinette et. al. (1993) have derived a verbal-nonverbal factor split in some
populations. Bornstein and Chelune (1989) reported that the factor structure is influenced
significantly by variables such as the age and education of the subjects.

The Recurring Figures Test (Kimura, 1963) is a visual recognition task consisting of
either complex geometric figures or nonsense images presented on cards. Initially 20 cards are
shown to the subject. Then the subject is shown a series of 140 cards for 3 seconds each. Eight
of the original 20 images repeat seven times each, for a total of 56 exposures to pre-exposed
images. The remaining 84 images are unique, seen only once. The task of the subject is to
discriminate previously viewed images from unique images by saying "yes" if they had seen the
image before, and "no" if they had not. When Kimura (1963) administered the test to patients
with right- and left- temporal lobectomies, she was able to discriminate between the two groups
based on a significantly higher false-positive rate for right temporal lobe patients.

The Continuous Visual Memory Test (CVMT) (Trahan and Larrabee, 1988) is very
similar to the Recurring Figures Test (Kimura, 1963) but is somewhat shorter. In addition to
the recognition format-discrimination task, a 30-minute delayed-recognition task has been
incorporated in which subjects must choose the recurring figure from an array containing 7 foils.
Although patients with both left or right lateralized strokes performed significantly below
normal controls, Trahan, Larrabee, and Quintana (1990) report that significantly more patients
with right hemisphere lesions (63 % on delay) fail the CVMT than those with left hemisphere
lesions (23%). Poor performance has also been noted in patients suffering from Alzheimer's
disease (Trahan and Larrabee, 1985). Hall, et al (1996) noted a significant age effect on
CVMT, with elderly normal subjects performing poorly on most measures. In fact, using the
cutoff scores in the manual as many as 62% of normal elderly subjects were misclassified as
showing memory impairment. The above-mentioned studies comprise the bulk of the limited
body of research performed investigating the CVMT.

While individual tests are often able to make discriminations based on laterality, factor
analyses of groups of memory tests don't always support the validity of the constructs of verbal
and nonverbal memory. Smith et. al. (1992) noted that factor analysis of batteries of memory
assessment devices has yielded inconsistent results, with some authors deriving verbal-nonverbal
factor splits (Milner, 1972; Bornstein and Chelune, 1988) and others failing to do so (Lee,
Loring, & Thompson, 1989). Robinette, Sherer, and Adams (1993) derived consistent verbal
and nonverbal factors when utilizing a number of memory tests. The verbal factor consisted of
stories and paired associates from the WMS and the Luria Memory Test at 30-second delay.
The nonverbal memory factor consisted of the Tactual Performance Test-Memory and Tactual
Performance Test-Location measures from the Halstead-Reitan battery, subtest #7 from the
Halstead-Reitan Category Test, and the immediate and delayed figural reproduction measures of
the WMS, as revised by Russell (1975). The measures were orthogonally rotated to isolate the
possible confound of verbal mediation of nonverbal material. The tests as a whole were able to
consistently differentiate normal subjects from those with documented, medical brain injury;
however, no information was gathered as to the laterality of damage in the brain injured
subjects. As noted earlier, factor analysis of the WMS-R has failed to consistently resolve
verbal and nonverbal memory factors.

The literature on hemispheric asymmetry supports the idea that verbal and nonverbal
processing are different, and that the two hemispheres show differing abilities to perform these
tasks. History and anecdotal information suggest that patients with right hemisphere lesions
tend to show deficits in visual/nonverbal processing and memory. The tests above have been
developed to assess these abilities. However, tests of nonverbal memory show inconsistent
performance in identifying a nonverbal memory factor, and sometimes fail to distinguish patients
with left versus right hemisphere lesions. The question we must ask ourselves, then, is "Why are our tests unable to consistently identify verbal and nonverbal memory factors, and why are we unable to consistently identify individuals with right hemisphere lesions?" One explanation is that the tests are not pure measures of nonverbal memory.

**Nonverbal memory confounded by verbal mediation**

Chelune and Bornstein (1988) in analyzing the WMS-R suggest that the alleged visual memory tests are confounded on several levels. Part of scoring of WMS-R Visual Reproductions, or any constructional task such as the Complex Figure or the BVRT involves determining the accuracy of the drawing. As a result of application of strict scoring criteria, poor performance may result from carelessness, tremor or other motor deficits, and impaired visual perception, any of which may be completely separate from memory processes.

Yet another possible confounding factor is subject age. Smith et. al. (1992) reported that in studies that yield verbal-nonverbal factor splits, subjects are usually younger or members of clinical populations. Studies using older or non-patient subjects often yield unitary memory factors with no observable verbal-nonverbal split.

An important criticism that is consistently reported is the potential for what may be termed verbal contamination of nonverbal tests (Lee, Loring, & Thompson, 1989). The vast majority of these tests utilize simple, common geometric shapes for construction. These shapes have names, which allows for verbal encoding and rehearsal. Additionally, Lezak (1995) points out that in addition to naming figures, relative locations and descriptions can be verbally encoded (small circle left of the big triangle), allowing for verbal contamination. Vanderplas and Garvin (1959) noted that even with randomly generated, novel, complex shapes, verbal associations might be made with as little as three seconds of exposure.

In the validation of the Recurring Figures Test, Kimura (1963) administered a number of additional tests by displaying images tachistoscopically on a projection screen. The images included overlapping common figures, overlapping nonsense figures, arrays of letters and arrays of dots. On the overlapping figures presentations, subjects were to point out on a card the
images displayed on the screen. For letters subjects were to report the letters on the screen, and for dots subjects were asked to report the number of dots on the screen. Additionally, a group of familiar objects were cut from magazines, pasted on cards, and photographed for tachistoscopic slide presentation. Subjects viewed the slides and pointed to the corresponding images on their cards. The results of the testing indicate that patients with right temporal lobectomies showed inferior performance to patients with left temporal lobectomies on Dots, Overlapping Nonsense Figures, and Recurring Figures (as described above). On Letters and Overlapping Familiar Figures, no significant group differences were observed, and right lesioned patients showed superior performance to left lesioned patients on individual Familiar Figures. Groups were then compared to a small group of frontal lobe patients on the three tests in which no significant differences were noted. No group differences were noted between either of the temporal lobe groups and the frontal lobe patients, suggesting that the scores were not due to equal levels of reduced functioning. In summary, the right temporal lobe patients performed worse only on unfamiliar, visually presented material.

One potential explanation of these findings is that familiar figures have readily accessible verbal identities or names, which are easily processed and rehearsed verbally in the left hemisphere. Patients with left hemisphere damage, on the other hand, can utilize the intact right hemisphere to process the visual stimuli and buttress the impaired verbal coding of familiar objects. This is particularly supported by the superior performance by right hemisphere patients in remembering the single familiar image, which can be described as the most verbally loaded image in the study, and thus the most sensitive to verbal processing deficits.

Using the BVRT, Arenberg (1977) investigated the effects of verbal augmentation of visual retention. Subjects were administered the BVRT, which contains multiple alternate forms. During Form C of the test, an audio tape was played which described the figures seen on each plate of the test, giving names and descriptions of size, location, and other image characteristics. Timing of administration was standard, with the only deviation being the addition of the verbal description of the images. Subjects were compared to control groups
Verbal Mediation of the CVMT

who were given standard administrations of the BVRT. Relative to control subjects, members of the experimental group made significantly fewer errors on form C than other forms of the test. This main effect was magnified with older subjects. The results indicate that verbal encoding of visual stimuli on the BVRT is not only possible but improves performance. Additionally, the greater effect in older subjects is of special interest. Older adults are characterized as being less likely to elaborately rehearse information and provide additional retrieval cues under normal circumstances, but are able to utilize them when made explicitly available (Arenberg, 1977). The age x condition interaction suggests that younger adults may engage in this verbal encoding spontaneously while older subjects need more prompting to make effective use of verbal information in visual items.

Also using the BVRT, Helmstaedter, Pohl, and Elger (1995) explicitly examined the verbal encodability of the visual stimuli. Verbal descriptions of the BVRT items were generated, with points assigned for each verbal descriptor. Descriptions were given for form, position, and size of each figure on a plate, with a total score range from 2 to 15 for each plate, yielding an estimate of each image's verbal load. Patients with unilateral temporal lobe epilepsy (TLE) of the right or left hemisphere were compared to normal controls on performance on the BVRT. Subjects were also given a verbal memory test. Verbal memory was significantly worse in left TLE patients than right TLE patients or controls. Visual Retention on the BVRT separated brain injury patients from controls, but did not discriminate between right and left hemisphere groups. However, a high correlation was found between verbal memory ability and BVRT performance in the right TLE group, but not left TLE patients. This suggests that verbal memory may have mediated performance in the right TLE group. Furthermore, loss of information for each item increased as the verbal load of the material increased. As verbal load increased, differentiation between right TLE and left TLE groups also increased, with right TLE patients performing worse. Significant differentiation did not occur between TLE groups, however, until the verbal load surpassed 9 points, the critical value for 7+/- 2 units in short-term memory capacity. Thus, when verbal load surpassed short-term verbal capacity, performance in
right TLE patients dropped drastically. Normal controls having both systems intact performed best, and left TLE patients performed less well, possibly due to reaching maximum load for visual memory and lacking verbal compensatory mechanisms.

Swanson (1983) examined verbal coding of novel, irregular geometric figures in normal, verbal learning-disabled, and deaf children. In the first experiment, subjects were randomly assigned to either the named or unnamed condition. In the named condition, subjects were exposed to a nonsense geometric figure which had a verbal label associated such that the name was, to some extent, a meaningful representation of the figure. Examples of names for these figures were tooth (#19), badge (#25), and helmet (#26). Subjects were then asked to recall the figures and names using a probe-type serial recall task. Subjects in the unnamed condition used identical stimuli and test procedures, but naming was omitted. A significant main effect was found for naming in the normal and deaf children, but not for the verbal learning-disabled sample. Thus normal and deaf children benefited from verbal cues, but children with verbal learning disabilities did not. This is consistent with verbal processing deficits in children with verbal learning disabilities. It is likely that these deficits inhibited the children's use of verbal information to cue retrieval of visual images.

In the second experiment, the same procedure was used using only verbally learning disabled (VLD) and normal groups, with the naming group subdivided into spontaneous naming and forced naming. In the forced naming condition, subjects used the same stimulus labels as in experiment 1. In the spontaneous naming condition, subjects generated their own labels for the nonsense figures. Results indicate a main effect for group membership, with controls in both naming conditions performing significantly better than VLD subjects in the same conditions. Results further indicate that no significant main effect existed for forced versus spontaneous naming, and no group x condition interactions were noted. This supports the finding from Experiment 1 that VLD children do not benefit from verbal encoding of visual stimuli. Experiment 3 investigated forced naming in younger and older children in VLD and normal groups, and found main effects only for age and group (normal or VLD), but not for naming.
However, a significant group by condition interaction was noted, and subsequent analysis indicated that this interaction masked a true main effect for naming condition, with normal subjects benefiting from naming and VLD subjects failing to benefit. The summary conclusion of the study is that children with verbal learning disabilities do not benefit from verbal coding of nonverbal stimuli while other groups (deaf children, normal controls) do. This supports the notion of verbal mediation of visual stimuli in individuals with normal functioning left hemispheres.

Nagae (1977) conducted a study with college undergraduates to examine the degree to which verbal and visual encoding strategies interact with complexity of visual stimuli. Subjects were given a recognition test for complex, irregular geometric shapes and assigned to either a relevant or irrelevant naming condition. Similar to Swanson (1983) the relevant naming condition related the label to characteristics of the abstract figure. In the irrelevant naming condition the verbal label did not represent the figure or its characteristics. Results indicated that recognition was greater in the relevant naming condition than in the irrelevant condition. Furthermore, this relationship only existed with more complex shapes. Nagae (1977) interpreted the data to suggest that with highly complex shapes, the visual system is overtaxed, and that a relevant verbal label helps to encode and retrieve the figural memory.

In summary, verbal mediation or "verbal contamination" of nonverbal tests appears to exist in several tests, and is consistent with current understanding of cerebral hemispheric dominance and asymmetry. The task for those attempting to develop a pure test of nonverbal memory is to create images and procedures that do not promote relevant labeling. The images must be complex and novel, to eliminate verbal familiarity with simple shapes. They must also be either simple enough that they do not overtax the visual processing system, setting up a pure right hemisphere task, or must be designed in such a way that verbal descriptions do not aid in discrimination. The tests used must also be free of confound by other cognitive or motor processes which may affect memory, such as visual discrimination and learning.
Of the tests described above, the CVMT seems to show great potential for avoiding verbal mediation of stimuli. The CVMT is the only commonly available nonverbal memory test that also provides normative data for acquisition and learning, and separates this construct from memory. Furthermore, unlike many other nonverbal memory tasks, the CVMT includes an explicit check for visual discrimination. This addresses the possible confound of inferior visual processing separate from the memory construct. The CVMT utilizes a recognition format, rather than the more common constructive or reproduction format, which introduces confounds of visual perception and motor control. Stimulus presentation is very brief (only 2 seconds per item), minimizing opportunity for verbal encoding and rehearsal, and stimulus items consist of complex, assumedly novel visual images. The CVMT also incorporates a delay procedure, which helps to differentiate deficits in learning and acquisition from memory deficits. Given these qualities, it addresses many confounds and criticisms of other nonverbal memory testing.

The test construction does not, however, avoid all possible confounds to nonverbal memory. Vanderplas and Garvin (1959) noted that subjects are able to attach meaningful verbal labels to similar stimuli with very brief stimulus exposure. Little research has been performed using or investigating the CVMT. Explicit testing of verbal mediation of CVMT items has not been performed, raising the question of “How pure a measure of nonverbal memory is this test?”

The current study is separated into two phases. The first experiment was designed to describe the verbal content of the images of the CVMT. It was hypothesized that groups of images in the CVMT will differ in their respective verbal loadings, and that individual items will vary in verbal loading. The items and item classes used in the CVMT represent different kinds of visual stimuli and differing levels of complexity. As each of these factors has been associated with verbal contamination in the past, it was believed that similar findings would be apparent when examining the CVMT. Verbal responses to stimulus items were recorded and used to calculate indices of verbal load for each item and class of items, with each index measuring a different verbal characteristic of the items.
Verbal Mediation of the CVMT

The second phase of the study investigated the possible relationship between verbal loading of CVMT items and subject performance on delayed recognition of these items. It was hypothesized that significant relationships would be found between verbal loading of items and subject performance on delayed recall. If verbal contamination has the effect of enhancing visual memory performance, higher verbal loading should relate to increased memory performance.

Additionally, performance at delayed recall was compared for subjects in the two experimental groups. It was hypothesized that the group encouraged to make verbal associations to items would correctly identify more target items at delayed recognition than control subjects. This is consistent with findings reported by Arenberg (1977) indicating superior performance on the BVRT by subjects who were encouraged to verbalize versus control subjects.

EXPERIMENT 1

Method

The goal of the first experiment was to establish values for verbal loading of each item from the CVMT. Verbal loading was described in terms of general associability, proportion of content-laden responses given to the items, and heterogeneity of response both for individual items and for groups or classes of items from the CVMT described below.

Participants

121 undergraduate students enrolled in an introductory psychology class participated as subjects in the first portion of the study. Students were awarded credit for participation, in partial fulfillment of requirements of the class. Subjects were screened for current and past neurological disorders, including seizures and head injuries, as well as drug use using a brief screening form (See Appendix A) and the SCL-90R, a checklist of common psychiatric and psychological symptoms. 70 subjects were rejected from the data analysis due to past or
present treatment for neurological, psychiatric or psychological disorders, self-report of recent, significant difficulty with sleep, memory, or attention (reported on the SCL-90R), or significant current or past drug use. The remaining 51 subjects were retained for the data analysis.

**Apparatus**

Subjects were also asked to fill out an informed consent sheet for audio taping, a brief screening form (see Appendix A) and the SCL-90-R. Stimulus items were images from the Continuous Visual Memory Test (Trahan and Larrabee 1988). Audio tapes were recorded using a small \( (4'' \times 9'') \) audiocassette recorder with an amplified microphone positioned no more than 2 feet from the subject.

**Procedure**

Subjects were greeted and seated in the experimental room. After reviewing and signing consent forms, including informed consent for audio taping, subjects were read scripted instructions (See Appendix D), adapted from Trahan and Larrabee (1988) and Vanderplas and Garvin (1959).

Subjects were asked to give verbal descriptions first, and then to assess whether items are new or old. This deviation from standardized administration occurred for two reasons. First, the associative evaluation is likely to precede the memory portion of the test in all subjects, so that the deviation is simply making manifest a process that may already be taking place. Second, the information of primary interest is the verbal associations made by each subject. The memory portion of the test is retained both for purposes of standardization as well as future analysis of memory ability and potential comparison to other groups.

Subjects were administered all 137 items from the CVMT. Items were presented in the order in which they appear in the test booklet, according to instructions in the test manual.
Verbal Mediation of the CVMT (Trahan & Larrabee, 1988). Test administration was audio taped, and content responses were transcribed from the tapes. Subject responses of “new” or “old” were recorded on the CVMT protocol by the item administrator. After completing the sample and acquisition items, subjects were allowed 30 minutes to fill out the screening device as well as the SCL-90-R. After the 30 minute delay, subjects were administered the delayed recognition and visual discrimination tasks from the CVMT. Subjects were then debriefed and excused.

Indices of Verbal Load

Four indices were used to measure verbal loading of items or item classes. Three of these have been adopted from previous research conducted by Vanderplas and Garvin (1959) and differ in the specificity of the verbal characteristics of individual items. The fourth index is new, and is a measure of the verbal characteristics of a class of items rather than individual stimuli.

Proportion of Verbal Associations

The proportion of associative responses, designated index A was calculated for each image and for each class. For each image, index A was calculated by totaling the number of responses that included “Yes” or a verbal label (also called content responses). This total was divided by the total number of subjects to yield a percentage of subjects who made verbal associations to each item.

For each class index A was calculated by totaling the number of associative responses each made to items within a stimulus class. This total was divided by the number of stimulus items within the class to yield a percentage of associative response for each subject, for each class. Because classes varied in the number of stimulus items contained by each, raw scores
would not allow easy comparisons between classes. Mean values were calculated across subjects for each class to derive the A index for each class.

**Proportion of Content Responses**

An index of content responses, designated index C, was also calculated for each item and each item class. For individual items, this index was calculated by summing the number of content responses given by subjects and dividing by the total number of subjects. In other words, only responses representing verbal labels, such as "a pile of sticks," were counted. "Yes," responses were not coded as content responses. The result is an index of percentage of subjects that gave verbal content responses for each item.

For item classes, the same procedure was used as for calculating index A, but only verbal content responses were counted for each subject, excluding responses of "yes".

Index C is essentially a subset of index A, and is designed to be a more refined measure of verbal association as it counts only those verbal labels expressed by the subjects.

**Heterogeneity of Response to Individual Items**

The third measure of verbal load is an index of heterogeneity of response for each stimulus item (H-item). For each item, content responses made by all subjects were separated into categories of response (see coding of responses, below). Each response was compared to previous content responses for a given item. If a response was defined as synonymous to a previous content response to the same item the responses were grouped as a category of response. Responses that did not match with prior response classes in the list were assigned a new category. This process was repeated for each response to each item, yielding an array of categories of response to each stimulus item. The probability of a response belonging to a category was calculated by summing the number of responses in a given category and dividing
Fig 1. Examples of stimulus items representing, from left to right: Top row: Solid, 7-point polygons; Hollow-10 point polygons; Solid, 12-point polygons; irregular nonsense figures. Bottom row: Complex line drawings; checkerboard patterns; random patterns of sticks.

by the total number of responses. This probability value was entered into the entropy formula for information of verbal responses developed by Shannon and Weaver (1949): 

\[ H = -(\text{Sum} \ P(i) \log P(i)) \]

where \( P(i) \) is equal to the probability of a response belonging to the \( i \)th category. This index of heterogeneity of response indicates the degree of similarity between the specific verbal labels assigned by subjects to a given item.

**Heterogeneity of Response Within Design Classes**

In addition to the \( H \)-item, an index of heterogeneity of response was calculated for each class of response (\( H \)-class). Trahan and Larrabee (1988) describe the CVMT items in the acquisition phase as belonging to seven categories or classes (see Fig. 1): solid, random 7-point polygons; hollow, random 10-point polygons; solid random 12 point polygons; irregular nonsense figures, complex line drawings, checkerboard patterns, and randomized patterns of sticks. For each subject, responses to each item were sorted into the appropriate class or
category. Then H-class was calculated for each subject, following the same procedure as for H-item. The H-class value describes the distinctness of items grouped in the same design class. In other words, H-class is a measure of how verbally similar or different items within an item class are. This is particularly important given the delayed recognition format of the CVMT where subjects must select a target item from an array of items from the same class.

Coding of items

The values of the H-item and H-class were dependent upon coding of responses into post-hoc categories. Coding was performed by trained undergraduate research assistants. To help achieve interrater reliability, a sample of 10 subjects were run as pilot training data. Research assistants met with the author as a group and coded the pilot data, as was later performed during the coding of research subjects. Response classification was based upon consensus of multiple raters in the training session. When unanimous consensus was reached for the coding of the pilot data, the coding of the research data was performed independently by the trained research assistants.

Results

Verbal Encoding of Individual Test Items

The information for individual items derived from the above scales is represented in Table 1. Target items are those items with repeated exposures during the test. H-item values were calculated only for the first and last administration. As a result, some items do not appear in the table, as they were repeats of previously displayed stimuli. For verbal load comparisons, only the H-item values from the first exposure were used to describe target items. Association values (index A) ranged from 8% to 66. Content responses (index C) showed a similar range, from 6% to 66%. Due to very few "Yes" responses given by subjects there is a high degree of
Table 1. **Association, Content, and Heterogeneity of Response (H-item) for Individual Items.**

<table>
<thead>
<tr>
<th>Item</th>
<th>Index A (%)</th>
<th>Index C (%)</th>
<th>H-item</th>
<th>Item</th>
<th>Index A (%)</th>
<th>Index C (%)</th>
<th>H-item</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>22%</td>
<td>22%</td>
<td>.911</td>
<td>58</td>
<td>34%</td>
<td>32%</td>
<td>0.413</td>
</tr>
<tr>
<td>2</td>
<td>62%</td>
<td>58%</td>
<td><strong>0.604</strong></td>
<td>60</td>
<td>28%</td>
<td>26%</td>
<td>1.068</td>
</tr>
<tr>
<td>3</td>
<td>18%</td>
<td>18%</td>
<td>.954</td>
<td>62</td>
<td>20%</td>
<td>20%</td>
<td>0.94</td>
</tr>
<tr>
<td>4</td>
<td>78%</td>
<td>76%</td>
<td><strong>0.966</strong></td>
<td>64</td>
<td>34%</td>
<td>30%</td>
<td>0.962</td>
</tr>
<tr>
<td>5</td>
<td>20%</td>
<td>18%</td>
<td>0.151</td>
<td>65</td>
<td>32%</td>
<td>32%</td>
<td>1.13</td>
</tr>
<tr>
<td>6</td>
<td>22%</td>
<td>18%</td>
<td>0.723</td>
<td>68</td>
<td>50%</td>
<td>46%</td>
<td>0.68</td>
</tr>
<tr>
<td>7</td>
<td>40%</td>
<td>40%</td>
<td><strong>1.211</strong></td>
<td>69</td>
<td>24%</td>
<td>20%</td>
<td>1</td>
</tr>
<tr>
<td>8</td>
<td>66%</td>
<td>66%</td>
<td><strong>0.85</strong></td>
<td>71</td>
<td>54%</td>
<td>50%</td>
<td>1.05</td>
</tr>
<tr>
<td>9</td>
<td>22%</td>
<td>22%</td>
<td>1.136</td>
<td>72</td>
<td>32%</td>
<td>26%</td>
<td>1.07</td>
</tr>
<tr>
<td>10</td>
<td>36%</td>
<td>30%</td>
<td>1.28</td>
<td>75</td>
<td>50%</td>
<td>38%</td>
<td>0.676</td>
</tr>
<tr>
<td>11</td>
<td>28%</td>
<td>24%</td>
<td><strong>1.029</strong></td>
<td>76</td>
<td>54%</td>
<td>52%</td>
<td>0.905</td>
</tr>
<tr>
<td>12</td>
<td>14%</td>
<td>14%</td>
<td>0.555</td>
<td>78</td>
<td>70%</td>
<td>58%</td>
<td>0.872</td>
</tr>
<tr>
<td>13</td>
<td>34%</td>
<td>34%</td>
<td><strong>1.042</strong></td>
<td>80</td>
<td>22%</td>
<td>16%</td>
<td>0.432</td>
</tr>
<tr>
<td>14</td>
<td>48%</td>
<td>48%</td>
<td>0.759</td>
<td>82</td>
<td>32%</td>
<td>28%</td>
<td>1.044</td>
</tr>
<tr>
<td>15</td>
<td>26%</td>
<td>26%</td>
<td><strong>0.958</strong></td>
<td>85</td>
<td>38%</td>
<td>30%</td>
<td>0.92</td>
</tr>
<tr>
<td>16</td>
<td>34%</td>
<td>34%</td>
<td>1.111</td>
<td>86</td>
<td>50%</td>
<td>48%</td>
<td>0.776</td>
</tr>
<tr>
<td>17</td>
<td>18%</td>
<td>16%</td>
<td>0.903</td>
<td>87</td>
<td>28%</td>
<td>26%</td>
<td>0.876</td>
</tr>
<tr>
<td>18</td>
<td>44%</td>
<td>42%</td>
<td>0.984</td>
<td>89</td>
<td>10%</td>
<td>6%</td>
<td>0.477</td>
</tr>
<tr>
<td>21</td>
<td>38%</td>
<td>34%</td>
<td>0.735</td>
<td>90</td>
<td>30%</td>
<td>24%</td>
<td>0.96</td>
</tr>
<tr>
<td>22</td>
<td>32%</td>
<td>32%</td>
<td>0.654</td>
<td>91</td>
<td>52%</td>
<td>44%</td>
<td>0.958</td>
</tr>
<tr>
<td>24</td>
<td>48%</td>
<td>44%</td>
<td>0.467</td>
<td>93</td>
<td>44%</td>
<td>38%</td>
<td>1.013</td>
</tr>
<tr>
<td>25</td>
<td>8%</td>
<td>6%</td>
<td>0.477</td>
<td>95</td>
<td>26%</td>
<td>24%</td>
<td>0.96</td>
</tr>
<tr>
<td>27</td>
<td>14%</td>
<td>12%</td>
<td>0.678</td>
<td>97</td>
<td>32%</td>
<td>26%</td>
<td><strong>0.882</strong></td>
</tr>
<tr>
<td>28</td>
<td>20%</td>
<td>18%</td>
<td>0.728</td>
<td>98</td>
<td>44%</td>
<td>38%</td>
<td>1</td>
</tr>
<tr>
<td>31</td>
<td>64%</td>
<td>58%</td>
<td>0.945</td>
<td>99</td>
<td>62%</td>
<td>58%</td>
<td><strong>0.940</strong></td>
</tr>
<tr>
<td>33</td>
<td>40%</td>
<td>38%</td>
<td>0.666</td>
<td>100</td>
<td>32%</td>
<td>26%</td>
<td>0.958</td>
</tr>
<tr>
<td>35</td>
<td>52%</td>
<td>48%</td>
<td>1.052</td>
<td>101</td>
<td>28%</td>
<td>22%</td>
<td>0.911</td>
</tr>
<tr>
<td>36</td>
<td>26%</td>
<td>20%</td>
<td>0.728</td>
<td><strong>102</strong></td>
<td>54%</td>
<td>46%</td>
<td><strong>0.744</strong></td>
</tr>
<tr>
<td>38</td>
<td>32%</td>
<td>30%</td>
<td>0.412</td>
<td>103</td>
<td>22%</td>
<td>22%</td>
<td>0.713</td>
</tr>
<tr>
<td>40</td>
<td>48%</td>
<td>46%</td>
<td>0.811</td>
<td><strong>104</strong></td>
<td>68%</td>
<td>58%</td>
<td><strong>0.332</strong></td>
</tr>
<tr>
<td>41</td>
<td>22%</td>
<td>20%</td>
<td>0.94</td>
<td>105</td>
<td>24%</td>
<td>22%</td>
<td>1.041</td>
</tr>
<tr>
<td>44</td>
<td>42%</td>
<td>38%</td>
<td>1.171</td>
<td>106</td>
<td>32%</td>
<td>30%</td>
<td>0.96</td>
</tr>
<tr>
<td>45</td>
<td>24%</td>
<td>20%</td>
<td>0.857</td>
<td><strong>107</strong></td>
<td>46%</td>
<td>42%</td>
<td><strong>0.693</strong></td>
</tr>
<tr>
<td>46</td>
<td>26%</td>
<td>24%</td>
<td>0.91</td>
<td>108</td>
<td>66%</td>
<td>62%</td>
<td>0.938</td>
</tr>
<tr>
<td>51</td>
<td>42%</td>
<td>40%</td>
<td>0.63</td>
<td>109</td>
<td>56%</td>
<td>50%</td>
<td>0.818</td>
</tr>
<tr>
<td>52</td>
<td>18%</td>
<td>14%</td>
<td>0.759</td>
<td><strong>110</strong></td>
<td>42%</td>
<td>34%</td>
<td><strong>0.581</strong></td>
</tr>
<tr>
<td>54</td>
<td>52%</td>
<td>44%</td>
<td>0.937</td>
<td><strong>111</strong></td>
<td>66%</td>
<td>62%</td>
<td><strong>0.337</strong></td>
</tr>
<tr>
<td>56</td>
<td>28%</td>
<td>28%</td>
<td>1.103</td>
<td>112</td>
<td>42%</td>
<td>40%</td>
<td>0.728</td>
</tr>
</tbody>
</table>

Note: Higher values indicate higher verbal load for Associative and Content percentage; higher values indicate lower verbal loading for index H-item. Underlined items in boldface are first appearances of target items. Italicized items in boldface are last appearances of target items. Items not reported in the table represent repeats of target items within the test.
similarity between index A and index C. This is consistent with the findings of Vanderplas and Garvin (1959), where great similarity between A and C values was also found. The actual A and C values were also very similar to those found by Vanderplas and Garvin (1959), meaning that subjects in the current study gave verbal, content responses to the target stimuli as frequently as subjects in the previous study. H-item values ranged from 1.171 (item 44), indicating a high degree of heterogeneity of response, to 0.151 (item 5), indicating very similar responses across subjects. However, the lowest H-item values are associated with low C-indices, indicating that these low heterogeneity values stem from a low base-rate of responding and limited numbers of content responses to classify.

**Verbal Encoding of Items Within Classes of Stimuli**

A comparison of the seven classes of response revealed significant differences in verbal loading across classes. One-way, repeated measures ANOVA’s were calculated for each verbal load index across classes. Significant group differences were observed for item classes using index A ($F(6,300) = 15.61, p < .05$). Post hoc pairwise comparisons were made using Tukey’s

**Table 2. Mean A Values and Grouping for Each of the 7 Item Classes**

<table>
<thead>
<tr>
<th>Item Class</th>
<th>Mean A index</th>
<th>Standard Deviation</th>
<th>Verbal Load</th>
</tr>
</thead>
<tbody>
<tr>
<td>Irregular Nonsense Figures</td>
<td>56.0*</td>
<td>37.0</td>
<td>High</td>
</tr>
<tr>
<td>Random Sticks</td>
<td>54.7*</td>
<td>39.0</td>
<td>High</td>
</tr>
<tr>
<td>Checkerboard Patterns</td>
<td>54.1*</td>
<td>40.0</td>
<td>High</td>
</tr>
<tr>
<td>Solid, 12-point Figures</td>
<td>35.8 b</td>
<td>31.1</td>
<td>Low</td>
</tr>
<tr>
<td>Solid, 7-point Figures</td>
<td>34.9 b</td>
<td>29.5</td>
<td>Low</td>
</tr>
<tr>
<td>Complex Line Drawings</td>
<td>28.9 b</td>
<td>30.2</td>
<td>Low</td>
</tr>
<tr>
<td>Hollow Geometric</td>
<td>24.7 b</td>
<td>29.8</td>
<td>Low</td>
</tr>
</tbody>
</table>

Note: Higher A values indicate higher verbal load for each item class as measured by general level of verbal association. All values with the same superscript are statistically similar, and statistically different from values with different superscript notation.
HSD. The results indicate that for index A, classes split into high and low verbal load groups (see Table 2).

Index C also showed a high degree of variability across groups ($F(6,300)=15.54, p<.05$) indicating that the frequency of specific verbal labels being assigned to items was higher in some groups than others. Again, Tukey’s HSD was used to make post hoc, pairwise comparisons between groups. Again, groups separated into two distinct groups (See Table 3). Of note is the fact that the groups were comprised of the same classes for both index A and index C.

Table 3. Mean C Values and Rankings for Each of the 7 Item Classes

<table>
<thead>
<tr>
<th>Item Class</th>
<th>Mean C</th>
<th>Standard Deviation</th>
<th>Verbal Load</th>
</tr>
</thead>
<tbody>
<tr>
<td>Checkerboard Patterns</td>
<td>51.7</td>
<td>40.6</td>
<td>High</td>
</tr>
<tr>
<td>Irregular Nonsense Figures</td>
<td>51.6</td>
<td>36.5</td>
<td>High</td>
</tr>
<tr>
<td>Random Sticks</td>
<td>50.1</td>
<td>39.2</td>
<td>High</td>
</tr>
<tr>
<td>Solid, 7-point Figures</td>
<td>31.8</td>
<td>29.5</td>
<td>Low</td>
</tr>
<tr>
<td>Solid, 12-point Figures</td>
<td>30.1</td>
<td>29.5</td>
<td>Low</td>
</tr>
<tr>
<td>Complex Line Drawings</td>
<td>24.9</td>
<td>29.5</td>
<td>Low</td>
</tr>
<tr>
<td>Hollow Geometric</td>
<td>22.0</td>
<td>28.7</td>
<td>Low</td>
</tr>
</tbody>
</table>

Note: Higher C values indicate higher verbal load for each item class as measured by content associations. All values with the same superscript are statistically similar, and statistically different from values with different superscript notation.

H-item values were calculated by sorting responses made by all subjects. As a result, individual H-item scores are not available for individual subjects. In other words, H-item scores are a measure of the similarity in verbal responses to a given item, between subjects. As a result, a within-subjects ANOVA was not appropriate to compare the H-item scores across classes. The group averages and measures of variability are presented in Table 4. Mean H-item values for classes were calculated by including only the first presentation of target items. Repeated exposures to target items would over represent the contribution of a single item to the class average. Furthermore, repeated exposure to target items may result in changes in verbal load.
Table 4. Mean H-item Values for Each of the 7 Item Classes

<table>
<thead>
<tr>
<th>Item Class</th>
<th>Mean H-item index</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hollow Geometric</td>
<td>0.6746</td>
<td>0.26</td>
</tr>
<tr>
<td>Checkerboard Patterns</td>
<td>0.6913</td>
<td>0.15</td>
</tr>
<tr>
<td>Solid, 12-point Figures</td>
<td>0.7728</td>
<td>0.21</td>
</tr>
<tr>
<td>Random Sticks</td>
<td>0.8117</td>
<td>0.10</td>
</tr>
<tr>
<td>Complex Line Drawings</td>
<td>0.9253</td>
<td>0.22</td>
</tr>
<tr>
<td>Irregular Nonsense Figures</td>
<td>0.9919</td>
<td>0.09</td>
</tr>
<tr>
<td>Solid, 7-point Figures</td>
<td>1.0029</td>
<td>0.14</td>
</tr>
</tbody>
</table>

Note: Lower H-item values indicate greater similarity of verbal responses across subjects for each class and higher values indicate more idiosyncratic responding across subjects.

for a given item, thus skewing the H-item scores for the class. As a result, the class H-item mean represents the heterogeneity of response to items within a class for the first presentation.

The lowest H-item average was associated with the Hollow Geometric class, indicating that responses across subjects were most similar to each other for items in this class. Examples of these responses are "hollow shape" and "broken glass," two of the most common responses. The highest H-item average belonged to 7-point, solid polygons indicating high variability in verbal labels assigned to items in this class. Subject responses for this class ranged from simply descriptive "blacked in shapes" to "pants" and "boomerang".

H-item values were not, however, consistent for target items from the beginning of the test to the end of the test. In all cases, H-item values dropped for target items by the last exposure during the acquisition phase of the test (see Table 5). The changes all indicate increasing similarity in verbal labeling of target items across subjects. Similar comparisons of index A and index C do not demonstrate this pattern. Some items showed increases in index A and index C values from beginning to end, while others showed decreases in index values. No systematic change in A or C was observed. Increasing similarity in responses does not,
Table 5. Change in H-item Values From First to Last Exposure

<table>
<thead>
<tr>
<th>Item Class</th>
<th>H-item First Exposure</th>
<th>H-item Final Exposure</th>
<th>Change</th>
</tr>
</thead>
<tbody>
<tr>
<td>Irregular Nonsense Figures</td>
<td>1 0.966</td>
<td>0.94</td>
<td>-0.026</td>
</tr>
<tr>
<td>Hollow Geometric Figures</td>
<td>2 1.029</td>
<td>0.882</td>
<td>-0.147</td>
</tr>
<tr>
<td>Checkerboard Patterns</td>
<td>3 0.604</td>
<td>0.337</td>
<td>-0.267</td>
</tr>
<tr>
<td>Solid 7-point Figures</td>
<td>4 1.042</td>
<td>0.693</td>
<td>-0.349</td>
</tr>
<tr>
<td>Complex Line Drawings</td>
<td>5 0.958</td>
<td>0.581</td>
<td>-0.377</td>
</tr>
<tr>
<td>Solid 12-point Figures</td>
<td>6 1.211</td>
<td>0.744</td>
<td>-0.467</td>
</tr>
<tr>
<td>Random Patterns of Sticks</td>
<td>7 0.85</td>
<td>0.332</td>
<td>-0.518</td>
</tr>
</tbody>
</table>

therefore, appear to be the result of fewer total associative or content responses, as A and C values did not show a consistent pattern of increasing or decreasing over the test. This suggests that repeated exposure to target items may have the effect of increasing similarity in response across subjects as measured by H-item. This change may be due to characteristics of the items that subjects become increasingly sensitive to over repeated exposures. In other word, items may “pull” for certain responses, and this pull becomes more pronounced with repeated exposures, resulting in greater similarity in response across subjects.

The H-class index is a measure of the verbal similarity of items within each class of response for each subject. This differs from the H-item value discussed above. H-item measures similarity in response to a single item across subjects. H-class, in contrast, measures the similarity in response to multiple items within a class for a given subject. As such, H-class values are not available for individual items. Instead, each of the 7 classes is described by the H-class average across all subjects (see Table 6).

H-class values also showed systematic variation across item classes ($F(6,300)=6.33$, $p<.05$), indicating that the verbal distinctness of items within a class varied across groups of stimuli. Tukey’s HSD indicated that item classes split into high and low groups for verbal load
Table 6. H-class Values and Rankings for Each of the 7 Item Classes

<table>
<thead>
<tr>
<th>Item Class</th>
<th>Mean H-class Value</th>
<th>Standard Deviation</th>
<th>Verbal Load</th>
</tr>
</thead>
<tbody>
<tr>
<td>Irregular Nonsense Figures</td>
<td>0.2525</td>
<td>0.24</td>
<td>High</td>
</tr>
<tr>
<td>Solid, 12 point Figures</td>
<td>0.2243</td>
<td>0.26</td>
<td>High</td>
</tr>
<tr>
<td>Solid, 7 point Figures</td>
<td>0.2193</td>
<td>0.24</td>
<td>High</td>
</tr>
<tr>
<td>Complex Line Drawings</td>
<td>0.2181</td>
<td>0.28</td>
<td>High</td>
</tr>
<tr>
<td>Hollow Geometric</td>
<td>0.1731</td>
<td>0.25</td>
<td>High</td>
</tr>
<tr>
<td>Checkerboard Patterns</td>
<td>0.1204</td>
<td>0.19</td>
<td>Low</td>
</tr>
<tr>
<td>Random Sticks</td>
<td>0.1011</td>
<td>0.17</td>
<td>Low</td>
</tr>
</tbody>
</table>

Note: Higher ranking indicates higher similarity in responses within subjects for each class. All values with the same superscript are statistically similar, and statistically different from values with different superscript notation.

as measured by H-class. The highest H-class average, indicating a low level of similarity in verbal labels given to items within the class, belonged to Irregular Nonsense Figures. Subjects gave verbal labels to these items which were most distinct, separating items within the class. Responses varied from “spaghetti” to “wavy lines” to “person holding something.” The class with the lowest H-class average, indicating verbal similarity of items within the class, was Random Patterns of Sticks. Subjects tended to respond to all items in this class similarly, assigning labels such as "a pile of sticks" or "a bunch of lines" to all items within the class. The class that generated the most verbal associations and content labels was Random Patterns of Sticks. This class, however, demonstrated the least variability in responses to items within the class by individual subjects, indicating that the responses given were generic and descriptive of the class as a whole rather than individual item characteristics.
EXPERIMENT 2

Method

In addition to documenting the verbal load characteristics of the test items, the present study investigated the relationship between verbal loading and performance at delayed recall on CVMT target items. It was hypothesized that subject performance at delayed recall would differ significantly across items, and that this difference in performance would be related to verbal loading of target items. The empirical findings from Experiment 1 indicate that the CVMT is susceptible to verbal contamination, and that items and item classes show significant differences in verbal loading. Consistent with prior research, it was expected that verbal contamination would result in enhanced performance on easier-to-verbalize items. Furthermore, it was hypothesized that subjects in experiment one of the study would correctly recognize more target items at delayed recall than control subjects from experiment two, due to explicit instruction to engage in verbal labeling of items during the acquisition phase. These findings would be consistent with findings by Arenberg (1977) using a different test of nonverbal memory.

Participants

Subjects included 110 introductory psychology students. These subjects were recruited and screened as in Experiment 1 above. Of those subjects to participate, 59 subjects were rejected from the data analysis due to reported history of neurological or psychiatric illness, significant current or historic drug use, or reports of recent, significant difficulty with attention, memory, or sleep. The remaining 51 subjects were included in the data analysis.
Apparatus

Stimulus items were images from the Continuous Visual Memory Test (Trahan and Larrabee (1988). Subjects also filled out a brief screening form (see Appendix A) and the SCL-90-R. Subject responses were not recorded on audio tape as in the first experiment.

Procedure

After being seated in the experimental room, subjects reviewed and signed informed consent forms for testing. Research assistants administered the CVMT to subjects according to the standard instructions, included in the manual (Trahan & Larrabee, 1988). Subjects were read instructions from a script (see Appendix E).

Subjects were administered all items from the CVMT, according to standardized instructions (Trahan & Larrabee, 1988). Subject responses of new or old were recorded on the CVMT protocol by the examiner. After completing the sample and acquisition items, subjects were allowed a 30 minute break. During this time they were asked to fill out the screening device as well as the SCL-90-R. After the 30 minute delay, subjects were administered the delayed recognition and visual discrimination tasks from the CVMT. The delayed recognition task requires subjects to select the one target that was repeatedly presented (the “old” item) from an array, including foil items that were presented only once (“new” items) during the previous section of the test. Subjects were then debriefed and excused.

Results

Effect of Verbal Loading on Performance At Delayed Recognition

An initial within subjects, binomial analysis of variance (ANOVA) was calculated comparing scores on each item at delayed recognition (see table 7). At delayed recognition,
Table 7. Average Correct Performance at Delay for Each Item Class

<table>
<thead>
<tr>
<th>Item Class</th>
<th>Mean Score</th>
<th>Standard Deviation</th>
<th>Grouping</th>
</tr>
</thead>
<tbody>
<tr>
<td>Irregular Nonsense Figures</td>
<td>0.92</td>
<td>0.27</td>
<td>High^a</td>
</tr>
<tr>
<td>Solid, 7-point Figures</td>
<td>0.90</td>
<td>0.30</td>
<td>High^a</td>
</tr>
<tr>
<td>Complex Line Drawings</td>
<td>0.86</td>
<td>0.35</td>
<td>High^a</td>
</tr>
<tr>
<td>Checkerboard Patterns</td>
<td>0.86</td>
<td>0.35</td>
<td>High^a</td>
</tr>
<tr>
<td>Solid, 12-point Figures</td>
<td>0.65</td>
<td>0.48</td>
<td>Low^b</td>
</tr>
<tr>
<td>Random Sticks</td>
<td>0.61</td>
<td>0.49</td>
<td>Low^b</td>
</tr>
<tr>
<td>Hollow Geometric</td>
<td>0.53</td>
<td>0.50</td>
<td>Low^b</td>
</tr>
</tbody>
</table>

Note. All values with the same superscript are statistically similar, and statistically different from values with different superscript notation.

each subject attempted to pick the target item from an array including same-class foils. Correct identification was scored as a 1, and incorrect identification a 0. The ANOVA indicated that, at delayed recognition, subjects were consistently able to identify some targets with greater accuracy than others (F(6,300)=8.89, p<.05). Pairwise comparisons using Tukey’s HSD indicated that classes split into high and low groups for subject performance at delayed recognition. This opened the possibility that differences in performance may relate to differences in verbal load of the target items. Because several indices of verbal load were used for each item or class of items, several comparisons were made.

For each measure of verbal load an analysis was performed to test the hypothesis that higher verbal load of items is related to increased subject performance at delayed recall, as measured by accurate identification of target items. Additionally, each set of comparisons was made for verbal characteristics of both item classes and individual items.

Comparisons for Individual Items

The verbal load indices for individual items were used for the first set of comparisons. Because the recognition task involves selecting a target item from an array of foils, the verbal contamination of the individual item may result in enhanced ability to recognize these items.
Figure 2. Graphs indicate percentage of subjects that correctly identified target items at delayed recognition for visual representation. Items are ordered according to level of verbal load for each index, 1=highest verbal load, 7=lowest verbal load. Letters indicate items that were statistically similar in level of performance.
Items were ranked according to verbal load indices and subject performance at delayed recognition was plotted as the dependent variable (see Figure 2).

To examine the existence of a pattern of relationship between verbal load and subject performance at delayed recognition, target items were assigned ranks according to verbal load, as measured by H-item, index A, and index C, and were arranged in order by increasing verbal load. Target items are those stimulus items that were exposed multiple times ("old" items) during the acquisition portion of the test. High ranking indicates higher levels of verbal contamination as measured by each index, while lower ranks indicate lower verbal loading for individual target items. The dependent measure, correct identification of targets at delayed recognition, was used to make comparisons. As displayed in figure 2, no consistent trend was manifest relating high verbal load of individual items to higher performance at delayed recognition.

Comparisons for Item Classes

Because the recognition task involves making decisions about all items in a class, it is possible that the verbal characteristics of the class itself may influence subject performance. Comparisons of performance based on class characteristics were made to examine this possibility. Instead of individual pairwise comparisons, comparisons were made using the verbal load groupings empirically derived in experiment one. For each index the top two classes in the high groups and the lowest two classes in the low groups were combined, resulting in mean performance scores for each subject according to high versus low verbal loading. Due to variability in performance within the high and low groupings, the comparison was made by combining more than one class as a representative. The low group for H-class consisted of only two classes. To make all comparisons similar, and to allow equal variance
Within high and low groups, two classes were selected as representatives for each test. By choosing the top and bottom 2 classes, the comparison approximates a quartile split, which was felt to be the most accurate representation of the high-low comparison. Paired t-tests were used to compare performance at high and low levels for subjects (see Figure 3).

For H-class the high group was represented by Irregular Nonsense Figures and Solid 12-point Figures, and the low group was represented by Checkerboard Patterns and Random Patterns of Sticks. The performance comparison did not indicate any significant differences between the high and low group at delayed recognition ($t(50)= .84, p>.05$).

For index A, the groups that represented the high verbal loading condition were Irregular Nonsense Figures and Random Patterns of Sticks. The low condition was represented by Complex Line Drawings and Hollow Geometric figures. The results indicate that there was no significant difference between high and low verbal loading groups at delay using index A ($t(50)= 1.12, p>.05$).

For index C the groups that represented the high group were Checkerboard Patterns and Irregular Nonsense Figures. The representatives of the low group were Complex Line Drawings and Hollow Geometric figures. The results indicate that subjects were more accurate in identifying targets for items in the high verbal load group than the low verbal load group ($t(50)=3.73, p<.05$).

**Verbal vs. Nonverbal coding task**

Total scores at delayed recall were calculated for subjects in the first phase of the study (verbalizers) and subjects in the second experiment (controls). Performance data was spoiled for 5 subjects, due to scoring errors, clerical errors, or missing scores, in the first experiment, leaving a remaining $N=46$. No significant difference in performance was observed between the
Figure 3. Graphs indicate percentage of subjects that correctly identified target items at delayed recognition for each measure of verbal load for item classes. High and low groupings are based on empirical groupings derived from Experiment 1.
two groups. Verbalizers correctly identified an average of 5.61 items at delayed recall, versus 5.27 items for control subjects. \((t(95)=-1.27, p>.05)\). This indicates that the process of verbalizing responses did not significantly improve subjects' ability to correctly identify targets at delayed recognition.

**Summary**

Preliminary analysis indicated that performance by subjects at delayed recognition differed significantly across the 7 test items. Comparisons were made between groups with subject performance used as the dependent measure. Results indicate that the verbal load of individual target items does not relate significantly to differences in subject performance at delayed recognition. However, differences in performance did appear to be related to the differences in verbal load of the classes of items, as measured by index C, but not index A or H-class.

**Discussion**

The purpose of this study was to examine the possibility of verbal mediation of nonverbal items in a relatively new, popular, and little researched test of visual memory. The results indicate that the items of the CVMT are susceptible to verbalization and that some items and item classes are more easily verbalized than others. The differences in verbal load of individual items does not appear to relate to subject performance at delay, but differences in the verbal load of the classes may relate to subject performance at delayed recognition.

Verbal contamination of the test items, while evident, was minimal as measured by the verbal load indices used. Generally, low verbal association values were found for CVMT items. On average only 34% of subjects made content associations with individual CVMT items. These are comparable to findings by Vanderplas and Garvin (1959). However, H-item values
were quite low compared to prior research. Vanderplas and Garvin (1959) obtained heterogeneity values ranging from 2.0 to 4.4 indicating little similarity in verbal labels across subjects, whereas the current study found a maximum H-item score of 1.14. This indicates that verbal labels attached to CVMT items are more similar across subjects than those attached to the items used in the previous study. One explanation for this may be the visual complexity of the items. Vanderplas and Garvin used polygons, similar to two classes of CVMT items, but of differing complexity. Less complex, 4- and 6-point polygons yielded the highest H values in their research (comparable to the present H-item), while more complex items yielded lower H values. Items in the current study tended to be more visually complex, particularly the random sticks and checkerboard patterns. In the current study, subjects tended to give brief, nonspecific descriptions of such items. This resulted in less heterogeneity of response for specific items lowering H-item values. This tendency is also reflected in the low H-class values across all classes. Subjects tended to label items within classes similarly, suggesting that they were making associations to the general type of image, rather than specific details of the images themselves.

One area of concern, however, is the increase in verbal contamination of target items, as demonstrated by decreasing H-item values, from the beginning to the end of the test. While the current study does not indicate that this has an impact on a subject's ability to recognize targets at delay, it should be noted that repeated exposure may increase verbal contamination. The rate of change in verbal load is an empirical question to be addressed. The question of how many exposures are necessary to effectively learn the items and how this relates to verbal contamination could be addressed, resulting in a clearer idea of the costs and benefits of
This information could be useful in the development of other tests of nonverbal memory.

The different classes of items differed in verbal load, indicating that some item classes lend themselves more to verbal encoding than others. The classes which generated the highest rates of verbal responding, as indicated by index A, were Irregular Nonsense Figures and Random Patterns of Sticks. These classes were also found to be in the high group for verbal load as measured by index C, the measure of content-laden verbal labels. This indicates that items in these groups stimulated the highest level of verbal labeling by subjects. These two groups, however, were in the low group for the H-class index. This indicates that while the rate of responding was high, the distinctness of the labels assigned to items within the class was quite low. Subjects were readily able to assign verbal labels such as “checkerboard” and “pile of straws” to items in these classes, but assigned the same generic labels to all items in the class.

In contrast, the classes of Hollow Geometric figures and Complex Line drawings comprised the low group for both index A and index C, indicating that they generated the fewest total verbal responses. These item classes were also found together in the high group for H-class, indicating that they generated distinct verbal responses. This inverse relationship between response rate and specificity of response is of particular interest. Fewer total responses in combination with higher specificity of verbal label suggests that subject variables may be important. It may be that a subset of subjects is more likely to make specific verbal associations to test items, resulting in high specificity when overall response rates are low. The relatively high standard deviations for H-class support this notion. Future research may include investigations of subject variables as they relate to verbal labeling. In general, however, it
seems reasonable that test items that generate low overall rates of verbal labeling would be most useful for testing visual memory. Future test developers may wish to use items similar to Hollow Geometric and Complex Line Drawing items, which generated the lowest rates of verbal association in the present study.

In addition to describing verbal characteristics of the CVMT items, an initial empirical test of the effects of verbal contamination at delayed recognition was conducted. The results indicate that no systematic relationship exists between verbal load and performance at delayed recognition for verbal load of individual items. For classes as a whole, however, there is evidence that verbal load may have an impact on subject performance. Subjects performed better at identifying targets at delay for item classes that were characterized by higher rates of content association than for item classes with low content load. The similarity of responses within the class did not relate significantly to performance at delay. This suggests that the contaminating factor is not the salience of the item within the class. Rather, it is the verbal load of the entire class that is important. For items such as Checkerboard Patterns, the high rate of content response may indicate a general familiarity with the item type. For example, the Checkerboard Patterns were easily recognizable as belonging to a class and the general form was similar for all items. Half or more of the subjects that made content responses used a variant of the label “Checkerboard”. This familiarity of pattern may have allowed subjects to attend selectively to the details that made one item different from another. This is consistent with findings by Kimura (1963) and Helmstaedter, Elger, and Pohl (1995), where familiar items showed more verbal contamination and higher rates in subject accuracy in recalling the items. This may lead future test developers to use difficult to verbalize item classes, such as Complex Line Drawings and Hollow Geometric figures to assess nonverbal memory.
The hypothesized superior performance at delay for verbalizers versus controls was not found in the current study. This suggests that verbal encoding does not confer the same advantage on the CVMT as has been observed with the BVRT (Arenberg, 1977) or other tests (Swanson, 1983). The comparison in performance by verbalizers versus controls may have been affected by a low ceiling effect or by a restriction in the range of possible scores. With only seven potential responses, the total scores for each subject were limited in range from 0 to 7. This limited range may have had the effect of reducing variability of scores, minimizing differences in performance between the two groups. The restriction in range is may be a potential flaw in the clinical utility of the test as well as in research, especially with older subjects as described by Hall, et. al (1996).

While the current study offers partial support to the hypothesis that verbal load of items positively affects performance, the question of verbal contamination has not been completely answered. More research of the verbal characteristics of the CVMT is warranted given the present findings. Success in correctly selecting the target from an array including similar foils for each class was quite high, with accuracy rates over 80% for 4 of the 7 targets. No target was correctly identified by fewer than 52% of subjects. This suggests a ceiling effect in performance given a healthy, normal population, reducing variance and potential reliance on verbal cueing. The subjects in the current experiment were carefully screened to rule out neurological insult or other factors which may reduce performance. Intact visual memory processing may have reduced reliance on verbal cueing, thus reducing the strength of relationship between verbal load and performance at recall. A population of clinical subjects who may be more prone to accommodate visual processing deficits through verbal mediation should be used to investigate the possibility of verbal contamination using the CVMT. Similar
studies conducted using other tests, including the Recurring Figures Test (Kimura, 1963) and the BVRT (Helmstaedter, Pohl, & Elger, 1995), have used clinical samples and found support for the verbal contamination construct.

Other research of interest in the area of verbal contamination is suggested by the structure of the CVMT. Delayed recognition is not the only measure of performance available on the CVMT. Hits and misses, as well as a ratio of correct and incorrect responses (d') are available for the acquisition phase. The possibility exists that verbal contamination has its effects on the CVMT during the learning or acquisition phase, as well as at delayed recall. The high level of performance by subjects, as noted above, may mask some effects of verbal mediation at delay. Verbal mediation or contamination may affect the number of trials needed to acquire the visual pattern of the target items. Similarly, repeated exposures appears to increase verbal load of some items. Using this approach the effects of verbal contamination on nonverbal learning can be examined, in addition to the effects on memory.

These questions are not only of interest to researchers and users of the CVMT, but to the developers of other nonverbal memory tests. As pointed out by Moye (1997), construct validity research examining the question of verbal contamination, among other variables, is necessary for the development and selection of a variety of reliable, valid measures of nonverbal memory. Item analyses examining the verbal characteristics of individual items, the issues of exposure time, response format, and the presence or absence of a delay period are all factors that have been identified as potential weak points in tests of nonverbal memory. The CVMT addressed many confounds and appears to be less susceptible to verbal contamination than other tests, but did so by making several changes at once. The active ingredient(s) in verbal contamination are still unknown. Does shorter exposure time result in less verbal
contamination? Does the delay period result in a degrading or enhancing of verbal labeling? Researchers and developers of future tests must investigate these questions to adequately address the confound of verbal contamination and to increase the reliability and validity of nonverbal memory tests.

One weakness of this area of research is the lack of a "gold standard" for measuring and defining verbal characteristics of nonverbal items. Reviewing the literature, no two studies investigating different tests used identical procedures. This makes it difficult to compare verbal loading of items across multiple tests. Future research should focus on using a common method, which would help to define the characteristics of items less susceptible to verbal contamination. The outcome of such a standard approach would be information that would allow us to refine our tests of nonverbal memory. While meta-analysis of nonverbal memory tests allows some cross-comparisons of tests and variables, the differences in methodology make such comparisons difficult, time consuming, and compromise the validity of the findings. The method adopted for the current study is probably most useful in comparing tests like the CVMT, which use abstract designs rather than common geometric shapes as stimulus items. Tests such as the Abstract Designs-Recognition/Frequency, Biber Figure Learning, and others reviewed by Moye (1997) use similar stimuli and may be candidates for study using the current method. Such comparisons may yield a richer understanding of verbal labeling and verbal contamination of nonverbal memory.

In summary, the CVMT addresses the major confounds of construction, learning, and visual discrimination that plague most tests of nonverbal memory. The results of the current study indicate that the items of the CVMT are susceptible to verbal labeling, and that the different item classes differ in their verbal loading. While not every index of verbal load was
related to subject performance, there is evidence that verbal contamination of item classes related to subject performance at delay. Continued study of this test is recommended, utilizing other populations to provide a stronger test of verbal contamination. Further study examining the role of verbal contamination in the learning phase of the CVMT is also recommended. Finally, continued investigation of the verbal contamination construct is recommended.

Identification of which factors (type of stimulus, exposure time, delayed memory testing, response format, etc.) reduce verbal loading is necessary both to understand the phenomenon and to develop other valid and reliable tests of nonverbal memory. Adoption of a standard method for measuring verbal load is recommended to increase the validity of comparisons of different tests of memory. Research in these areas should result in both increased understanding of the construct of nonverbal memory and increased clinical utility of nonverbal memory tests.
Appendix A

Neuromedical Screening Form

ALL INFORMATION YOU PROVIDE WILL BE HELD STRICTLY CONFIDENTIAL
Please fill out this medical history questionnaire. When finished, place this form back into the envelope and seal it before returning to the examiner.

Neurological History:
1. Have you ever been evaluated or treated by a neurologist or neurosurgeon? ________
   If yes, please list condition: _____________________________________________

2. Have you ever had an injury to the head in which you received a concussion? ________
   If yes, how many concussions have you had? _________

3. Have you ever had an injury to your head that resulted in unconsciousness? ________
   If yes, how many times? _________
   For each instance, how long were you unconscious? _________________________

4. Have you ever had any seizures? ________

Psychiatric History
1. Have you ever been diagnosed with depression or any other psychiatric condition? ________
   If yes, please list diagnosis: ___________________________________________

2. Have you ever been hospitalized for mental health treatment? ________
   If yes, please list diagnosis: ___________________________________________

Drug History
1. Are you currently taking any of the following types of medication?: antidepressants, ________
   anticonvulsants (i.e., seizure medication), or tranquilizers?

2. Have you used hallucinogens or opiates more than 50 times? ________
   (e.g., LSD, Mescaline, Peyote, STP, DMT, Psilocybin (mushrooms),
   Heroin, Morphine, Opium)

3. Have you used marijuana or hashish in the past 24 hours? ________
   Have you used marijuana or hashish more than 4 times per week over the last year? ________

4. Have you used cocaine, crack, or ecstasy more than 50 times? ________

5. Have you used inhalants (e.g., glue, gasoline, nitrous oxide) more than 10 times? ________

6. Have you used stimulants (e.g., amphetamine) more than 20 times per year? ________

7. Have you used antianxiety agents or sleeping medication in the last 24 hours? ________

8. Have you used prescription pain medication in the past 24 hours? ________

9. Have you ever been treated for alcoholism? ________

10. Are you taking any medications not listed above at this time? ________

If yes, please list: ____________________________________________________


Appendix B

Informed Consent to Participate in Research

The purpose of this project is to study the characteristics of a test of memory. You will be asked to view drawings used in this test, and to try to remember which items repeat themselves. In addition, you will be asked to describe what, if anything, the drawings resemble to you. In addition to the testing, you will be asked questions about your drug use and medical history.

Your responses will be kept strictly confidential, and your name will not be associated with your answers at any time. Only a subject number will be included on scoring sheets and screening instruments. Your responses will be recorded on audiotape, which will be transcribed later. At that time, the recording will be destroyed.

The test you will take is designed so that no one performs perfectly on it. You should not expect to complete all tasks without errors. It will take approximately one hour to complete the testing. You will be given a rest break approximately halfway through the session. At the completion of the testing, you will be awarded 2 experimental credits. The person that is conducting the testing will not be able to give you extensive feedback regarding your performance.

The benefit from this research is to allow students to participate in psychological research and to develop information about a commonly used test. There are no other alternative procedures which would produce these benefits. You are free to ask questions at any time during or after the testing. You are also free to withdraw your participation at any time, without penalty. Please feel free to ask any questions at such a time. If you should have any questions after you complete the experiment, please contact Sean Whalen at the Clinical Psychology Center, 243-4523, or Dr. Stuart Hall at his office, 243-5667. Although there is little risk of danger or harm to subjects, the following statement is required for informed consent:

In the event that you are physically injured as a result of this research, you should individually seek appropriate medical treatment. If the injury is caused by the negligence of the University of Montana or any of its employees you may be entitled to reimbursement or compensation pursuant to the Comprehensive State Insurance Plan, established by the Department of Administration under the authority of M.C.A., Title 2, Chapter 9. In the event of a claim for such physical injury, further information may be obtained from the University's Claims Representative or University Legal Counsel.

(Reviewed by University Legal Counsel, July 6, 1993)

By signing below you agree that you have read and understand this document. You agree to participate in this project with knowledge of potential risks and benefits.

Signature of Subject: ___________________________ Date: __________
Appendix C

Debriefing

The test you have just taken is a measure of nonverbal memory. The construct of nonverbal memory is useful in both clinical and research work in neuropsychology. The idea of nonverbal memory has been around for quite a while, and has been consistently used in clinical and research work, but has been questioned in empirical studies investigating memory testing. One potential confound of nonverbal memory is verbal mediation or processing of nonverbal stimuli. In other words, we sometimes may use words or labels to help us remember difficult visual items like the ones on the test.

The purpose of this study is to examine the potential verbal contamination of the test you have just taken, and to look at how this verbal contamination may affect subjects' performance. The results will have potential impact on interpretation of performance on this test, and may influence development of future tests of nonverbal memory.
Appendix D

Instruction Script, Experiment 1

"This is a measure of visual memory. During the next few minutes I will show you a series of designs, one at a time. Each card you see in the sequence will have one design drawn on it, and you will have 2 seconds to view it. Some of the shapes may remind you of some familiar object or situation, while others may not remind you of anything. Your job will be to name whatever the shape reminds you of, if anything. Some of the shapes may remind you of some object or situation, but you may not be able to describe it in the short time during which you see the shapes. If the shape reminds you of something that you can describe in a word or two, simply say that word or phrase. If you cannot describe it in a word or two, simply say yes. If the shape does not remind you of any object or shape, say "No." It is important that you say something, either a word or phrase, the word "Yes," or the word "No," for each drawing. Some of the designs in the stack will occur only once; so after you've seen them the first time they will not be presented again. However, a number of designs will be repeated numerous times throughout the series. Try to remember each design you see, so that if it is repeated later you'll be able to recognize it as one you've already seen. For each card, after trying to describe what it might remind you of, I would like you to think about whether or not you have seen the design before during this test. If you're seeing the design for the first time, it will be a new design to you because you haven't seen it before, so I'd like you to also say "New." However, if it is a design that is repeating itself in the deck and this is the second, third or fourth time you've seen it, it will be an old design to you, and so I'd like for you to say "Old." So for each item you will try to describe what it reminds you of, using a few words or by saying "Yes" or "No." You will also tell me, for each item, whether it is a new or old design. Do you understand"
Appendix E

Instructions Script, Experiment 2

"This task is a measure of visual memory. During the next few minutes, I will show you a series of designs, one at a time. Each card you will see in the sequence will have one design drawn on it, and you will have 2 seconds to view it. Some of the designs in the stack will occur only once; so after you've seen them the first time, they will not be presented again. However, a number of designs will be repeated numerous times throughout the series. Your task will be to try to remember each design you see, so that if it is repeated later, you'll recognize it as one you've already seen. As we go through this deck of cards, I'd like you to look at each design. If you're seeing the design for the very first time, it will be a new design to you because you haven't seen it before, so I'd like for you to say "New." However, if it's a design that is repeating itself in the deck and this is the second, third or fourth time you've seen it, it will be an old design to you, and so I'd like for you to say "Old." Do you understand?" (Trahan and Larrabee, 1988 p. 6)
Verbal Mediation of the CVMT

References


