VALIDATION OF A NEUROPSYCHOLOGICAL WADA PROCEDURE

APPROVED BY SUPERVISORY COMMITTEE

This work is dedicated to my parents, who instilled in me their values and traditions, with unconditional love and support.

VALIDATION OF A NEUROPSYCHOLOGICAL WADA PROCEDURE

By

Daniel David Eisenman

DISSERTATION

Presented to the faculty of the Graduate School of Biomedical Sciences

The University of Texas Southwestern Medical Center at Dallas

In Partial Fulfillment of the Requirements

For the Degree of

DOCTOR OF PHILOSOPHY

The University of Texas Southwestern Medical Center at Dallas

Dallas, Texas

August, 2005

ACKNOWLEDGEMENTS

I would like to acknowledge the help and support of many people who contributed to this dissertation and my development as a psychologist. First of all, I would like to express my overwhelming gratitude and appreciation to Dr. Laura Lacritz, chairman of my dissertation committee and longtime mentor. As a mentor, Dr. Lacritz challenged me to think critically and demonstrated a commitment to high standards and scholarship that I will strive to maintain throughout my career. Dr. Lacritz's door was always open for questions and advice regarding the project. Her suggestions improved many aspects of my dissertation, and her careful reviews of my written work helped me express my thoughts with increased clarity. Her guidance throughout the project ensured that I was fully prepared for each stage of the endeavor. Dr. Lacritz has played a special role in my professional development, overseeing my graduate work in neuropsychology from beginning to end; she has always offered encouragement and support, without ever doubting my abilities.

I would also like to thank Dr. Munro Cullum whose comments and suggestions challenged me and taught me to think critically. His thorough reviews of my manuscript were an important part of my research development and significantly improved the quality of my written work. I would also like to thank Dr. Paul Van Ness who allowed me the opportunity to validate this IAP measure and set aside time to work with me as needed. Dr. Van Ness' expertise in epilepsy and enthusiasm for research in this area made the project more enjoyable. I would also like to thank Dr. Shannon Whyte for her open door policy, willingness to discuss research complexities, and her positive support and encouragement throughout the project. In addition, I would like to thank Dr. Carroll Hughes for his valuable feedback and support. I want to acknowledge the help of Dr. Linda Hynan who set aside time to help me with this study, Dr. Alan Frol who offered encouragement and agreed to meet with me to discuss aspects of the project, and Dr. Kathleen Saine who piqued my interest in epilepsy. I would like to acknowledge Dr. Carlos Marquez de la Plata who helped me run several statistics and Dixie Woolston who helped me format the final copy of this document. Special thanks to my parents for standing by me and helping me set my goals high, without wavering in the belief that I could reach them. I would like to thank my three sisters Lisa, Shira and Debbie as well as my brother-in-law Dr. Matthew Goldberg, for all their love and support. In addition, I would like to thank all my good friends for standing by me and offering encouragement. Last but not least, I would like to thank Liam Zaidel who was always available to problem solve obstacles and offer support.

VALIDATION OF A NEUROPSYCHOLOGICAL WADA PROCEDURE

Publication No. _____________

Daniel David Eisenman, Ph.D.

The University of Texas Southwestern Medical Center at Dallas, 2005

Supervising Professor: Laura Lacritz, Ph.D.

ABSTRACT

The Intracarotid amobarbital procedure (IAP) is considered an essential part of the presurgical neurodiagnostic evaluation in most epilepsy centers throughout the country. Despite the IAP's mainstream use, there is great variability in how the test is being performed across centers. The main purpose of this study was to validate a standardized IAP memory measure and explore its value in predicting lateralization as well as treatment outcome. The IAP memory measure had good overall reliability for all three Forms (Cronbach's Alpha = .85

for Form I, .83 for Form II, and .69 for Form III). The majority of items on Form I and II had acceptable item difficulty values, item discrimination values, and item-total correlations. There were a few items on Forms I and II that may be candidates for revision, but most items were only slightly below predicted ranges for what defines a "good" measure, and the majority of items contributed to the reliability of the test. In terms of construct validity, correlations with other memory tests provided some evidence of convergent validity for the IAP memory measure, but were generally low. In terms of divergent validity, both Forms I and II had low or no correlations with the executive functioning measures, providing preliminary support for the construct validity of the measure. When utilizing a discrepancy score to predict lateralization in subjects with temporal lobe epilepsy, asymmetry scores from the IAP memory measure were able to classify 92% of subjects with either left or right TLE after applying a correction factor for left injection scores. There was limited data regarding post-operative seizure outcome. However, seizure-free subjects had a higher percentage of DS greater than 20% than those subjects who reported at least one seizure postoperatively. In addition, there was a significant difference on Form I of the IAP memory measure between subjects who reported they were seizure free following temporal resection as compared to subjects who continued to report symptoms. Overall, the IAP memory measure demonstrated good psychometric properties and this study represents one of the most thorough analyses of the IAP memory test to date.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF ABBREVIATIONS

Electroencephalography .... (EEG)

Intracarotid Amobarbital Procedure .... (IAP)

Temporal lobe epilepsy ....(TLE)

Mesial temporal lobe epilepsy ....(MTLE)

Antiepileptic drug ....(AED)

Epilepsy monitoring unit ....(EMU)

Positron emission tomography ....(PET)

Electroconvulsive therapy .... (ECT)

University of Texas Southwestern Medical Center ..... (UTSWMC)

Magnetic Resonance Imaging .... (MRI)

Weschler Adult Intelligence Scale-Third Edition .... (WAIS-III)

California Verbal Learning Test .... (CVLT)

Weschler Memory Scale-Third Edition ....(WMS-III)

Discrepancy Scores ....(DS)

Full Scale IQ .... (FSIQ)

Right TLE ....(RTLE)

Left TLE .... (LTLE)

Wisconsin Card Sorting Test ....(WCST)

Rey-Osterrieth Complex Figure Test .... (Rey-O)

Anterior temporal lobectomy .... (ATL)

Chapter One: Introduction

A seizure is an excessive and abnormal electrical discharge of neurons in the brain evidenced by behavioral or Electroencephalography (EEG) changes, or both (Guberman, 1994). The term "epilepsy" encompasses a number of different syndromes manifested by recurrent unprovoked seizures. Epilepsy is one of the most common neurological problems worldwide. Approximately two million people in the United States suffer from epilepsy, and three percent of people in the general population will have epilepsy at some point in their lives (Wyllie, 2001). Epilepsy syndromes fall into two main categories: generalized and focal syndromes. In generalized epilepsies, the predominant type of seizure begins simultaneously in both cerebral hemispheres. Many forms of generalized epilepsy have a strong genetic component and in most, there is no distinct lesion. Generalized motor seizures are often characterized by loss of consciousness, generalized tonic-clonic movements of the limbs, tongue biting, and incontinence. In focal epilepsy, seizures originate in one or more localized foci, although they can sometimes spread throughout the entire brain. Focal seizures are classified as either simple (without alteration of consciousness) or complex (with an alteration of consciousness; Sadock & Sadock, 2003). Most focal epilepsies are believed to be the result of one or more central nervous system insults, but in many cases the nature of the insult is not identified (Chang & Lowenstein, 2003). Approximately 50% of individuals with epilepsy have focal epilepsy (Guberman, 1994).

Approximately 20% of individuals with epilepsy have seizures that cannot be controlled by antiepileptic drugs. It has been estimated that half of these individuals are potential candidates for surgical treatment (Engel & Shewmon, 1993). Correct lateralization and localization of epileptogenic origin is critical to the success of resective surgery in order to minimize damage to nearby regions of the brain. To assist in seizure localization, an extensive neurodiagnostic evaluation involving EEG, structural and functional neuroimaging techniques, the Intracarotid Amobarbital Procedure (IAP), and neuropsychological assessment is conducted on potential surgical individuals (Rausch et al., 1993).

The IAP is commonly included as part of the evaluation process and was first proposed in 1949 by Dr. Juhn Wada as a method to determine hemispheric lateralization of language functions (Wada, 1949). This technique involves the injection of sodium amobarbital through a catheter into the internal carotid artery. The ipsilateral hemisphere of the brain becomes anesthetized, leaving the hemisphere contralateral to the injection to function in isolation. Impairment of speech functions during this process became the “gold standard” for determining language dominance.

Milner, Branch, and Rasmussen (1962) extended the use of the IAP to include assessment of hemispheric memory competence prior to performing temporal lobectomy for treatment of intractable temporal lobe epilepsy (TLE). To assess memory performance during IAP, a subject is typically asked to recall items presented when the functions of one hemisphere are suppressed by the amobarbital. Use of the IAP for this purpose was further prompted by reports of global amnesia following unilateral temporal lobectomy (Corkin, 1984). The IAP was soon utilized by epilepsy centers across the country as a procedure useful in isolating the memory capabilities of each hemisphere and helping neurologists identify individuals with severe memory impairment contralateral to the site of surgical resection (Simkins-Bullock, 2000).

Use of the IAP to assess memory capabilities is quite common. A survey of 53 epilepsy surgery centers found that 98% use the IAP to assess hemispheric language dominance and memory functions prior to surgical resection of the temporal lobe (Rausch et al., 1993). Memory assessment during the IAP typically involves the presentation of a combination of pictures, objects, or words during the period of hemispheric anesthetization followed by recall or recognition of the information previously presented. Despite its widespread use, there is not a widely standardized administration procedure or means of evaluating memory function. Trenerry and Loring (1995) hypothesized that there are as many variations in IAP memory testing procedures as there are centers, making it difficult to compare outcome data across settings.

Multiple sources of variation in IAP memory testing protocols exist. There are differences in the manner in which the actual IAP is administered, including amount of drug injected, type of drug utilized, and rate of drug administration.

Substantial variability also exists in memory testing procedures administered during the IAP. Centers vary widely in terms of type and number of memory stimuli presented, timing of stimulus presentation, method of memory testing (e.g., free recall versus forced-choice recognition), timing of recognition testing, and the definition of a “passing” or “failing” performance (Rausch et al., 1993). Important patient management decisions are made on the basis of IAP memory testing results, including the nature and extent of temporal lobe resection and, in some cases, whether surgery is performed at all. It is clear that current IAP memory procedures vary widely by evaluating different types of memory (recall versus recognition) at different points in time (during anesthetization of post-drug effects), and with different methods of interpreting the resultant data. These discrepancies pose limitations to widespread interpretation and application of IAP results, and highlight the need for a standardized method of testing memory functions during the IAP.

Chapter Two: Temporal Lobe Epilepsy

TLE is a type of focal epilepsy defined as recurrent seizures most often originating from mesial temporal lobe structures, although seizures can also arise from lateral temporal lobe structures. Approximately 80% of complex partial seizures arise from the temporal lobes, and TLE is the most common form of epilepsy in adults, affecting approximately 3 in 1,000 people (Sadock & Sadock, 2003). Of the 1,200,000 Americans with focal epilepsy, TLE occurs in over 400,000. The onset of TLE is often in early childhood, but in approximately one third of the patients, seizures typically disappear for several years, returning in the teens or early adulthood (Kasper, 2005). A history of febrile convulsions is found in approximately 40% of cases of intractable TLE (Devinsky, 2004). TLE can be subclassified into mesial temporal lobe epilepsy (MTLE) and lateral temporal neocortical epilepsy. This subclassification is useful, because MTLE comprises the majority of the cases of epilepsy refractory to pharmacotherapy (Babb & Brown, 1987).

Typical mesial temporal lobe seizures usually begin with auras such as a sensation of epigastric rising, emotional features (most commonly fear), and occasionally psychic symptoms or olfactory or gustatory hallucinations, all indicating involvement of mesial temporal limbic structures (Guberman & Bruni, 1999). Once the seizure progresses, individuals frequently lose awareness and stare blankly. Automatisms or nonpurposeful movements are common including chewing, lip smacking, or swallowing, or more elaborate behaviors such as a display of extreme emotion or running. This phase lasts a minute or so, during which the patient is amnestic and apparently unaware of the environment. Then there is a subtle transition to increased environmental responsiveness, with more coordinated and purposeful activity, including intelligible verbalization. The ictal phase usually lasts 2 or 3 minutes and rarely more than 5 minutes (Davenport, 1998). Afterwards there is confusion for varying periods of time, and the individual may experience some postictal aphasia if the seizures began in the language-dominant hemisphere. (Chang & Lowenstein, 2003).

Individuals with TLE often show impairments in attention, memory, processing speed, executive functions, mood, and personality. The most common interictal cognitive complaint is impaired memory. In terms of emotional symptoms, interictal depression occurs in approximately one third of TLE individuals. Nearly 20% of individuals with TLE report suicidal ideation and suicide rates in TLE are more than 10 times higher than in the general population (Devinsky, 2004).

TLE is treated with medications, vagus nerve stimulation, and resective surgery. Approximately one-third of patients with epilepsy do not respond to treatment with a single antiepileptic drug, and it becomes necessary to try a combination of drugs to control seizures (Kasper, 2005). An estimated 90% of hippocampal atrophy cases are intractable. Patients with MTLE have a much lower response to medication and an estimated 75% of patients with MTLE fail most forms of medical treatment aside from surgical resection (Spencer, 2002). If adverse effects develop from one antiepileptic drug (AED), another monotherapy is usually tried. When seizures persist despite high concentrations of AEDs in the blood, a trial of two AEDs is sometimes recommended. When monotherapy fails, two AEDs will improve seizure control in more than one third of individuals, albeit with only 10% experiencing full control of seizures (Devinsky, 2004).

The vagus nerve stimulator (VNS) is a relatively new nonintracranial surgical approach to the treatment of epilepsy. Vagus nerve stimulation involves the implantation of an electric device that stimulates the vagus nerve and thus reduces seizure activity. The vagus nerve is one of the cranial nerves and has motor functions in the larynx (voice box), autonomic functions in the gastrointestinal system and heart, and sensory functions in the ears and tongue. Stimulation of the vagus nerve is thought to affect connections to areas in the brain that are prone to seizure activity although its precise mechanism of action is not entirely clear. Individuals with complex partial seizures who do not respond to anticonvulsant medication and who are not good candidates for resective surgery are often considered for vagus nerve stimulation therapy (Devinsky, 2004). Approximately one-third to one-half of individuals implanted with VNS have greater than 50% seizure reduction when used in addition to medication therapy (Guberman & Bruni, 1999). However, VNS rarely provides full seizure control and medication reductions are modest when VNS is used as adjunctive therapy (Devinsky, 2004).

Individuals with temporal lobe seizures who are nonresponsive to medication are diagnosed with refractory epilepsy (Engel, 1996). Epilepsy surgery is frequently considered in these individuals since approximately 60% to 80% of individuals are rendered seizure-free after temporal lobectomy (Devinsky, 2004). Correct lateralization and localization of epileptogenic origin is critical to the success of resective surgery in order to minimize damage to nearby regions of the brain. Accurate localization is often difficult and it is estimated that only 10% to 15% of individuals with refractory partial seizures undergo surgical resection, in part due to the difficulty in localizing the seizure focus (Devous, 1995).

To facilitate correct localization, an extensive neurodiagnostic evaluation is conducted to help determine the seizure focus. EEG, the IAP, brain imaging, and neuropsychological testing are common parts of a comprehensive assessment of surgical candidates. Twenty-four hour EEG and video monitoring is common for surgical candidates and usually takes place on an epilepsy monitoring unit (EMU). It is often necessary to reduce AEDs to precipitate seizures during this period of monitoring, and some centers reduce medication abruptly while others tailor AED treatment to each individual patient based on seizure frequency. (Rausch et al., 1993).

Neuroimaging is an initial part of the neurodiagnostic evaluation since it is useful in detecting the presence of hippocampal sclerosis, a common lesion found in many individuals with TLE. In hippocampal sclerosis, there is selective loss of neurons in the hippocampus. More specifically, the CA2 region is relatively spared in terms of cell loss, while approximately 75% of CA1 cells and 50% of CA3 neurons die (Wieser et al., 1993). The dense gliosis that accompanies the loss of neurons causes shrinkage and hardening of tissue. The term "mesial temporal sclerosis" has also been used for this lesion, because often there is neuronal loss in the neighboring entorhinal cortex and amygdala (Chang & Lowenstein, 2003). Although structural changes are concentrated in the hippocampus, amygdala, and adjacent limbic cortex, functional disturbances are more extensive. Hypometabolism on positron emission tomography (PET) in TLE patients often involves the entire temporal lobe as well as the ipsilateral thalamus (Duchowny, Harvey, Sperling, & Williamson, 1998). There continues to be great debate and uncertainty as to whether hippocampal sclerosis is a cause or an effect of seizures (Chang & Lowenstein, 2003). Detection of mesial temporal sclerosis is important, since 80% of individuals with this lesion can expect to become free of disabling seizures following anterior temporal lobectomy (Guberman & Bruni, 1999).

The neurodiagnostic evaluation often involves neuropsychological evaluation, which consists of using standardized, measures to characterize patients' neurocognitive functioning. While the specific neuropsychological measures used in the evaluation vary from center to center, the overall test battery typically encompasses a common set of neuropsychological abilities. Among these cognitive domains are measures of global intelligence, attention and concentration, learning and episodic memory, language, fine motor skills, and executive functioning (e.g., problem solving, mental flexibility, abstract reasoning; Trenerry, 1996). Neuropsychological assessment is currently the best available means of quantitatively describing a patient's cognitive status.

Memory testing is an important part of neuropsychological testing in individuals with epilepsy (Rausch, Le, & Langfitt, 1997). Memory tests traditionally include measures that involve the recollection of verbal (e.g., word lists or stories) or visual information (e.g., geometric designs). The neuropsychological evaluation provides a baseline assessment of cognitive abilities prior to surgery and allows the opportunity for individuals to be retested following surgery as a means of comparing pre and post-surgical cognitive functioning (Rausch et al., 1993).

In addition, neuropsychological testing can yield diagnostic information that aids determining the location or lateralization of the seizure focus (Snyder & Nussbaum, 1998). More specifically, left hemisphere lesions, particularly those involving the left temporal lobe, tend to result in lower performance on tasks involving verbal learning and recall. Similarly, lesions in the right hemisphere are often associated with lower performance on learning and recall of spatial or nonverbal information. Consequently, neuropsychological profiles offer useful information which can aid in determining lateralization of seizure focus.

Neuropsychological evaluations are also frequently used to estimate neuropsychological outcome following surgical resection particularly with respect to memory functions. Investigators have long sought methods for identifying individuals who are most at risk for developing memory complications following surgery. Chelune (1995) explored two hypotheses regarding memory functioning: hippocampal adequacy versus functional reserve. The functional reserve model asserts that the contralateral temporal lobe to resection is more predictive of patients at risk for developing post surgical memory deficits. Therefore, if temporal resection is to occur on the left side, then poor IAP memory scores on the right side would signify the likelihood of memory deficits. The hippocampal adequacy model states that the functioning of the temporal lobe where resection is to occur is more predictive of patients at risk for developing relative memory losses (Chelune, 1995). There is evidence from baseline neuropsychological studies and IAP investigations that the functional adequacy of the tissue to be resected in the ipsilateral temporal lobe determines the nature and extent of memory loss following surgery. While there is debate as to which hypothesis is more predictive of memory outcome following surgical resection, many neuropsychologists consider both theories when interpreting results.

Chapter Three: Intracarotid Amobarbital Procedure (IAP)

The IAP or Wada has become the "gold standard" for assessing hemispheric language dominance (Alpherts, Vermeulen, & van Veelen, 2000). The procedure involves injecting sodium amobarbital, a barbiturate, into one of the carotid arteries through a catheter placed in the femoral artery. The barbiturate temporarily anesthetizes the portion of the ipsilateral cerebral hemisphere perfused by the anterior and middle cerebral arteries for approximately five minutes. While one hemisphere is inactivated, language functions such as confrontation naming, repetition, and reading are assessed (Simkins-Bullock, 2000).

The first attempts to anesthetize the brain were made in the early 1940's when a Canadian neurologist named Juhn A. Wada developed a technique to inactivate certain brain regions (Snyder & Harris, 1997). Ironically, Wada's development of the IAP technique was somewhat fortuitous, as his primary aim was to use sodium amobarbital as a means to improve efficacy of electroconvulsive therapy (ECT) for treating psychiatric individuals with depression or psychosis. His goal was to use the IAP to determine speech dominance in order to place ECT electrodes closer to the non-speech dominant hemisphere.

In 1962, Brenda Milner and colleagues extended the use of the IAP to predict postsurgical amnesia. Milner originally advocated for the use of memory assessment during the IAP to determine whether significant mesial temporal lobe damage contralateral to the proposed resection was present. Her rationale was that after amobarbital injection there should not be any memory deficit unless there is a lesion in the hemisphere contralateral to injection. However if there is a lesion, then the temporary inactivation of the "normal" hemisphere by the amobarbital should transiently produce memory deficits characteristic of individuals with a hippocampal lesion (Milner et al., 1962). Milner and colleagues examined the use of the IAP for memory assessment in 50 individuals undergoing the IAP for speech lateralization. They found significant memory impairment (on recall of pictures of common objects presented prior to injection) in 11/50 individuals who were injected contralateral to the side of the seizure focus (Milner, Branch, & Rasmussen, 1962). Milner later replicated her findings in 226 individuals undergoing the IAP. She found a significant relationship between errors in recalling the pictures of objects and the side of pre-existing lesion (Milner, 1972). Milner's studies were the first to demonstrate how the IAP might be used to assess memory in individuals with a pre-existing lesion.

The IAP is a popular procedure and continues to be used throughout epilepsy centers around the world. At the 1992 Palm Desert Epilepsy Surgery Conference, 71 epilepsy surgery centers participated in a survey in order to gather more information about IAP's uses and procedures (Rausch et al., 1993). Of the 71 surgery centers, 68 reported using the IAP in epilepsy surgery candidates. The three not performing the IAP included a center in Russia, one in Italy and a new epilepsy center in the United States. Survey responses reported over 1,569 standard IAP procedures being performed each year. This figure is an obvious underestimate since there were many epilepsy centers who did not participate in the survey. The majority of epilepsy centers reported using the IAP for assessing both language and memory in temporal lobe surgery candidates (Rausch et al., 1993).

Certain aspects of the IAP appear common to most epilepsy centers (Acharya & Dinner, 1997). In most centers, a cerebral angiogram is performed shortly before the IAP to evaluate the vascular anatomy and to assess the degree of hemispheric cross-flow. This is important in order to document any anomalous patterns that may influence the distribution of the drug or affect the interpretation of results. The angiogram is also used to detect basilar artery filling which could cause the patient to lose consciousness and stop breathing. Following the angiogram a catheter is inserted in the internal carotid artery via the femoral artery. The patient is then asked to keep his/her arms elevated and is told to start counting aloud. Sodium amytal is then injected over several seconds. The contralateral arm drops as hemiplegia quickly develops. If the dominant hemisphere is injected, global aphasia typically develops and the patient becomes mute for a brief period (usually 2-3 minutes). Speech then returns but paraphasic errors are common for the next few minutes. In addition to aphasia, confusion and inattention can result from dominant hemisphere injection. When the nondominant hemisphere is injected, the patient continues to count but usually has some degree of dysarthria. Muteness may also occur but lasts only 20-30 seconds. Language and memory items are presented for the duration of the hemiparesis. In some epilepsy centers, both retrograde and anterograde memory functions are assessed. Anterograde memory assessment involves recall of items presented during the peak unilateral drug effect. Retrograde memory assessment refers to recall of items (after amobarbital has dissipated) presented to the patient prior to injection. The 1992 Palm Desert Survey indicated that 79% of participants "always to almost always" inject both hemispheres when evaluating memory competence (Rausch et al., 1993). In most epilepsy centers, the side of the presumed seizure focus is injected first, followed by the other hemisphere (Acharya & Dinner, 1997).

For the past 25 years, the IAP has been used to predict postsurgical memory outcome in epilepsy centers throughout the country (Acharya & Dinner, 1997). Despite its mainstream use, no standardized protocols exist. A recent review article regarding the variability of the IAP stated:

> Other than the vessel selected, there is very little that is "standard" about the IAP. Instead the IAP involves varying procedures, varying stimuli, varying personnel, and essentially differs to a lesser or greater extent in every institution employing the procedure. The drug dosages; the manner

> in which the drug is delivered; the method for determining or defining when the drug has, in fact, taken place; and the type of stimuli used; the timing of stimuli; and the criteria for "passing" and "failing" all vary from center to center (Simkins-Bullock, 2000, p. 42).

This makes for a difficult comparison of IAP data across testing centers and highlights a need for a more standardized procedure for the IAP.

In addition to differences in how the procedure is performed, there is also great variability in the behavioral measures of the IAP. In terms of stimuli, there are two procedures that are commonly used in order to assess memory. The first involves presenting several different stimuli to the patient. The patient is then presented a distractor task, followed by the presentation of the original stimuli again to see if he/she can recognize it. This format often continues until the patient's performance returns to baseline levels. The second procedure involves the presentation of a preset number of discrete stimuli during the period of the drug effect. Then the patient's recall or recognition memory for the stimuli is tested once the drug anesthetization has worn off (Loring et al., 1992).

Different test stimuli are used across centers, including tangible objects, line drawings of objects or designs, cards with printed words or sentences, spoken phrases or words, and photographs. Some centers recommend using stimuli that can be dually encoded (based on verbal and visual characteristics) as a means of identifying individuals at risk for post-surgical amnesia. Both verbal and nonverbal stimuli are recommended since reduced performance on verbal memory items has been demonstrated following amobarbital injection into the dominant language hemisphere (Loring et al., 1992; Loring et al., 1993). At the University of Texas Southwestern Medical Center (UTSWMC), the memory measure alternates between written words and pictorial items presented following injection of amobarbital.

Some studies have reported that certain test stimuli (e.g., real objects versus line drawings or objects versus words) may be more effective in assessing a patient's lateralized memory performance. One study conducted at a hospital in Sweden found that memory for pictures of common objects or for concrete words was sensitive to both left and right temporal lobe lesions while abstract words are predominantly processed by the intact left hemisphere. The authors reported that abstract figures and faces are processed by both hemispheres (Christianson, Saisa, & Silfvenius, 1990). Loring and colleagues (1997) found that real objects were superior to line drawings on the IAP in demonstrating lateralized temporal impairment.

There have been few studies comparing different IAP behavioral procedures in terms of prediction of surgical outcome. In one of the few studies comparing IAP procedures, Dodrill and Ojemann (1997) compared the percentage of agreement for three different methods of memory assessment for the IAP (in terms of “passing” the IAP versus “failing” the IAP) as well as the ability of the three methods to predict verbal memory deficits following epilepsy surgery.

The first memory procedure, known as “the Montreal” procedure (Branch, Milner, & Rasmussen, 1964) involves the presentation of five line drawings after the amobarbital has taken effect, followed by recognition memory assessment when the amobarbital has dissipated. The second procedure, referred to as “the Seattle” procedure has been used at the Washington School of Medicine’s epilepsy center for more than 20 years and is described by Dodrill and Ojemann (1997). The procedure involves the continuous presentation of a three part task: *Object naming* in which a solid object is presented to the patient followed by *Reading* where a short sentence printed on a card for the patient to read as a distractor, and *Recall* in which the patient is asked to recall the object presented to him prior to reading the card.

The last procedure, known as the Interview procedure, has been described by Dodrill and Ojemann (1997) and involves waiting until the amobarbital has dissipated and the patient has returned to “normal.” Then individuals are evaluated for five basic events that occurred during the period of drug effect. The events are: presence versus absence of speech blockage, presence or absence of motor weakness (the correct arm must be indicated based on side of injection), presence of absence of object naming errors, presence or absence of errors when reading sentences, and whether or not errors were made in recalling the items presented during the Seattle procedure (Dodrill & Ojemann, 1997). Passing the IAP following ipsilateral injection was defined as greater than three out of five correct for the Montreal procedure, less than 50% error rate for the Seattle procedure, and four out of five items correct for the Interview procedure.

Dodrill and Ojemann administered all three memory procedures to 172 individuals, with the Montreal items interspersed among the Seattle procedure at specified intervals followed by the Interview procedure once the drug effect had dissipated. The results indicated significant differences in the percentage of individuals passing across all three procedures. Passing of all three amobarbital procedures occurred in only 30% of the 318 injections. Individuals more frequently passed the Montreal and Seattle procedures than the Interview procedures. Dodrill and Ojemann hypothesized that the Interview procedure was the most diverse procedure, and unlike the other two, does not limit itself to recall of specific stimuli. The Montreal and the Seattle procedures were in the most agreement, with nearly 70% of cases in agreement with respect to passing and failing the IAP. The Seattle and Interview procedures had the least concordance, with agreement in only 45% of the cases (Dodrill & Ojemann, 1997). This study suggests that the same patient could have differing results in terms of passing or failing the IAP, depending on which epilepsy center and procedure is being utilized.

In order to explore the ability of the procedures to predict post surgical verbal memory deficits, the authors conducted postoperative memory testing on 60 of the 172 individuals involved in this study. They used both objective and subjective criteria to define post surgical memory deficits. The subjective criteria were simply the patient's self-report on whether they had experienced any memory loss. The objective criteria were defined as a 40% decline on two out of three standardized neuropsychological measures (Wechsler Memory Scale-Revised: Logical Memory-immediate recall, Logical Memory-delayed recall, and Paired Associates-delayed recall). Out of the 60 individuals, 15 met objective and subjective criteria for post surgical memory loss. The Seattle procedure offered the best prediction of post surgical memory deficits (based on subtests from the Wechsler Memory Scale), with 76% of cases correctly classified. The Interview identified post surgical verbal memory deficits correctly in 52% of the individuals, while the Montreal correctly predicted memory deficits in 48% of the individuals. The Seattle procedure had the lowest error rate and the highest sensitivity and specificity when compared to the other two procedures (Dodrill & Ojemann, 1997). There are some limitations in this study (e.g., using just the three subtests from the WMS-R as the only objective outcome measure); however, this study is the only reported attempt to compare different IAPs in their ability to predict memory deficits following surgery.

Timing of memory stimuli presentation also varies from center to center. The IAP memory testing protocol at the Medical College in Georgia involves the presentation of one set of stimuli at an "early" stage (before speech has recovered) and another set of stimuli at a "late" stage after language functions have returned (Loring, Lee, Meador, Flanigin, & et al., 1990). Other criteria used to determine timing of presentation of stimuli include a predetermined time limit after injection ranging from 0 to 5 minutes or the presence of hemiparesis with or without marked ipsilateral EEG slowing (Acharya & Dinner, 1997). The Palm Desert Survey indicated that 60% of respondents presented items within the first three minutes following amobarbital injection while the rest waited until after three minutes before presentation of memory items (Rausch et al., 1993). Timing of stimulus presentation is important because following amobarbital injection there is an initial period of muteness, confusion, disorientation, inattention, and impaired responsiveness especially when the language dominant hemisphere is injected. Lesser, Dinner, Luders, and Morris (1986) explored whether stimuli presented during the initial stage of confusion can be recalled after a delay. Lesser and colleagues investigated 36 individuals who had left injections and were left hemisphere dominant for language. The authors were investigating whether the initial muteness and apparent confusion after amobarbital injection prohibited the formation of new memories. For the purposes of the study, "confusion" was defined as a period of time occurring after cessation of counting and during which the patient was unable to maintain attention on a single object or person, except transiently. Immediately after the arm contralateral to injection fell, the patient was presented several different objects. Often it was necessary to hold a patient's eye open to show an object. After confusion began to clear EEG slowing began to dissipate, the patients were shown additional stimuli including words, phrases, and pictures. As soon as EEG slowing was no longer present and strength had returned to baseline the patient's memory was tested for the "early" objects presented immediately following injection as well as the "late" stimuli. Foils were used to control for confabulation. The results demonstrated that 18 out of 24 patients with left TLE and 4 out of 12 with right TLE recognized at least two-thirds of the "early" objects during post-testing. The authors concluded that the initial stage of confusion after amobarbital injection did not prevent the formation of new memories (Lesser et al., 1986).

To explore stimulus timing effects on IAP memory testing, Loring and colleagues at the Medical College of Georgia (1994) compared item presentation at three different time periods. Eight common objects were presented during the "early" period, which occurred about 30 to 45 seconds following injection. Five stimuli (two objects, one rhyme, and two designs) were presented during the "middle" phase, which occurred about 3 to 3 1/2 minutes following injection. Four stimuli (line drawings) were presented during the "late" phase, which occurred 5-1/2 to 6-1/2 minutes following injection. Results indicated that the memory performance on the early items was significantly better at predicting seizure laterality than was performance on the middle or late items (Loring et al., 1994). Consequently, it would be worthwhile for other epilepsy centers with different protocols to perform similar studies on items presented shortly after injection of amobarbital versus items presented several minutes following injection.

Another difference between epilepsy centers in terms of memory protocols is the number of test stimuli presented. Loring and colleagues present 8 tangible objects during their "early" stage, followed by presentation of 2 objects, a nursery rhyme, and 2 visual discrimination items during the "late" stage. The IAP protocol at the Cleveland Clinic involves presenting a maximum of 20 items, including 4 objects, 2 line drawings, 5 written words, 1 color, 1 number, 3 pictures, and 4 spoken phrases (Wyllie et al., 1991). The Seattle procedure mentioned earlier involves the continuous presentation of stimuli over a 5-minute period. The same 20 stimuli and distractors are used repeatedly throughout the duration of the IAP (Dodrill & Ojemann, 1997).

Memory for stimuli presented during the IAP is tested by free recall, forced-choice recognition, or multiple-choice recognition procedures. When recognition procedures are employed, the number of foils (i.e., distractor items) presented varies from center to center. Loring and colleagues (1990) utilize two foils for each target item, while the Cleveland Clinic's protocol includes three foils per target item (Wyllie et al., 1991). At the University of Texas Southwestern Medical Center (UTSWMC), the recognition component of the IAP memory protocol includes a larger number of foils, with 7 distractor items per target stimulus. This version was compared to another 16-item IAP protocol that utilized a forced-choice recognition format with no foils, showing that the 7-foil procedure correctly identified lateralized memory dysfunction at a higher rate (51% versus 29%). This suggests that a greater number of foils may enhance accuracy of the recognition component of IAP memory assessment (Epker et al., 2000).

Most centers complete the memory portion of the IAP when effects of the drug are believed to have resolved. Criteria for determining when drug effects have dissipated also vary, and include predetermined time intervals, resolution of hemiparesis, and return to EEG baseline (Acharya & Dinner, 1997). The procedure utilized by Dodrill and colleagues at the University of Washington differs in that recall is continuously assessed during the drug effect, rather than after the patient has recovered (Dodrill & Ojemann, 1997).

While most epilepsy centers have a method for determining whether a subject "passed" or "failed" memory testing following anesthetization of each hemisphere, there is considerable variability across epilepsy centers in how IAP data are interpreted. For example, clinicians at the Cleveland Clinic utilize a comparative method whereby memory results from each hemisphere are compared. Failure is defined by less than 67% retention on either side (Wyllie et al., 1991; Kneebone, Chelune, Dinner, Naugle, & Awad, 1995). The Seattle IAP method derives a memory score for the percentage of errors made within the five minutes of memory testing. Error rates of 0 to 49% are considered passing, while an error rate of greater than 50% constitutes failure (Dodrill & Ojemann, 1997). Loring and colleagues do not utilize a fixed pass/fail criterion for memory performance. They do require recognition of at least 2/8 memory stimuli in order to not repeat the Wada memory assessment, and are more comfortable with scores of at least 3/8 correct (Trenerry & Loring, 1995).

One procedure that has been described in several articles published by Loring and colleagues is the IAP used at the Medical College of Georgia. The procedure involves the presentation of eight common objects for 4-8 seconds each. The object names are repeated twice to the patient and include a combination of ordinary household items (e.g., fork, mousetrap), small toys (e.g., troll), and plastic food (e.g., hotdog, pizza). Recognition memory of the material presented during the procedure is tested after amobarbital effects have worn off with each of the 8 objects presented randomly interspersed with 16 foils. The patient is instructed to choose the items presented previously. To correct for response bias and guessing, Loring and colleagues incorporated a correction by subtracting one-half of the incorrect (false-positive) responses from the total number of correct responses (Loring et al., 1993).

Utilizing this procedure, Loring and colleagues (1993) reported a significant correlation (r = 0.78) between IAP memory asymmetries (differences in IAP memory scores between the left versus right hemisphere) and Magnetic Resonance Imaging (MRI) hippocampal volume asymmetries in 20 individuals with complex partial seizures using this procedure. In several other studies out of the Medical College of Georgia, scores from this protocol have been shown to be related to hippocampal pathology, seizure laterality, seizure outcome after surgical resection, and post-operative material-specific memory decline (Loring et al., 1997; Loring et al., 1990; Loring, Murro, Meador, Lee, & et al., 1993; Loring et al., 1994).

Due to the varied IAP protocols, there is a significant lack of reliability and validity studies. The 1992 Palm Desert Survey indicated that the lack of validation studies was a major concern of participants (Rausch et al., 1993). Most studies on IAP reliability involve looking at the results from individuals who were readministered the IAP to see if they had the same pass-fail classification. In the Palm Desert Survey (Rausch et al., 1993), more than a quarter of respondents reported they perform repeat IAPs in all of their individuals who fail their initial IAP. There are a few cases in the literature of individuals who failed an IAP, were retested, and then passed the second procedure. In a brief report, Dinner, Luder, Morris, Wyllie and Kramer (1987) readministered the IAP to five individuals and found improvement in their memory scores following the second administration. The authors speculated that the improvement may have been due (in part) to practice effects. In contrast to the speculations of Dinner and colleagues (1987), McGlone and MacDonald (1989) collected data on 70 patients, most of whom were epilepsy surgery candidates. The two hemispheres were injected on different days, in most cases 24 hours apart. The side injected first was counterbalanced across sex and hand preference. The order of injection (ipsilateral vs. contralateral) was not counterbalanced. The results indicated that recognition of memory scores was better following injection ipsilateral to the seizure focus as opposed to injection contralateral to the seizure focus. Further, they found no significant difference whether the right or left hemisphere was injected first. The authors concluded there was no evidence for a practice effect or transfer of learning (McGlone & MacDonald, 1989).

Jones Gotman (1992) examined 14 repeated cases at the Montreal Neurological Institute and Hospital. The interval between first administration and second administration ranged from 3 days to 14 years. Eleven of the 14 cases had the same results as on the first administration, with three cases being classified as failing after passing the first administration. She concluded that their IAP procedure is reliable since 79% of the individuals had the same results as their first IAP.

There is a need for individual centers to measure, monitor, and publish their success rates for predicting postsurgical memory changes so that centers with weaker reliabilities and validities can switch to methods with empirical support (Simkins-Bullock, 2000). In short, there is a clear need for an extensive investigation of the reliability and validity of different IAP methods in predicting postsurgical memory deficits.

The Palm Desert Survey (Rausch et al., (1993) referenced cases related to the validity of the IAP. Participant findings were divided into True Positives, False Negatives and False Positives. True Positives were defined as cases where individuals failed the IAP with injection ipsilateral to proposed resection and later developed post-surgical memory impairments; False Positives referred to individuals who failed the IAP but went on to benefit from surgical resection; and False Negatives were instances where individuals "passed" the IAP but later went on to develop postoperative amnesia or other memory complications. Of the 71 survey respondents, six indicated they "had knowledge of a patient who failed the memory component of the IAP (with injection ipsilateral to proposed surgery) and subsequently became amnesic (permanently or transiently) after temporal lobe surgery." Loring and colleagues (1994) described a case of post surgical amnesia in which the patient had a right seizure focus and recalled none of the 8 items presented after right hemisphere injection. The patient recalled all the items presented to him after left injection. This asymmetry or difference between right and left hemisphere performance is predictive of an individual at risk for postoperative amnesia. Other factors including neuropsychological testing predicted no increased risk of amnesia. The patient underwent right temporal lobectomy and following surgery, the patient developed a permanent anterograde amnestic syndrome.

Jones-Gotman (1992) reported on 11 out of 72 individuals who had failed their ipsilateral IAP. The 11 individuals underwent limited surgical resection. The postoperative performances were compared to the 61 individuals who passed their ipsilateral IAPs utilizing three subtests from the WMS-R (Logical Memory, Associative Learning, and Visual Reproduction). The results indicated that the 11 individuals who had failed the IAP performed significantly worse postoperatively than the 61 individuals who passed their ipsilateral IAP. There was some improvement in performance for the 11 individuals who failed the IAP after the one year follow-up. However, they still performed worse than the other 61 individuals who passed the IAP. The results speak to the validity of the IAP and the authors state the results "strongly support its [the Wada] validity as a screen to protect memory" (Jones-Gotman, 1992).

One of the purposes of the IAP is to act as a screen and prevent individuals from undergoing surgery that might result in severe postsurgical amnesia. However, there is some discussion in the literature regarding the IAP's specificity. In other words, is the IAP preventing individuals from having surgery who might have gone on to benefit from surgical resection?

Rausch and colleagues (1993) reported that the possibility of failing more individuals than necessary “is a major concern to most centers.” Stanulis and colleagues (1990) reported that only 4/180 consecutive individuals performed poorly enough on the IAP to require retesting and in three of the four cases, repeated IAPs resulted in a passing performance. All four cases underwent surgical resection without major complication. The authors stated that current pass/fail criteria are a misuse of the IAP, resulting in an unacceptably high rate of false-positive results and denial of modification of operative treatment” (Stanulis, Valentine, & DeToledo, 1990). Novelly and Wiliamson (1989) investigated 325 IAP protocols and reported that IAP failure was found in 25/325 (7.5%) of the cases. When the individuals underwent repeat testing using a lower dose of amytal, 21 out of 25 of them “passed” the second IAP and later underwent surgical resection without postoperative amnesia. These authors concluded that false positives can result from IAP and that retesting can be done with valid results.

Most of the research on the validity of the IAP is limited by the obvious challenges in conducting experimental research to assess the reliability and validity of the technique. It would be unethical to conduct a prospective study randomizing individuals into groups using the IAP as an independent variable. The literature is lacking in any comprehensive comparison of postoperative memory outcomes among individuals who did or did not undergo IAP testing. In addition, several epilepsy centers exclude data from individuals who do not undergo surgical resection because of results from IAP testing (Martin & Grote, 2002). Consequently, there is not a clear understanding of the reliability and validity of the IAP in predicting risk for postsurgical memory loss.

One way to gather more information regarding the reliability and validity of the IAP is for each epilepsy center to conduct an examination of their own IAP and investigate how well their procedure predicts different postsurgical outcomes. To better understand the IAP, each center needs to carefully examine the psychometric properties of their memory assessment procedures, and maintain a detailed database on individuals undergoing the IAP.

As discussed earlier, the IAP currently has two major uses: language/speech lateralization and prediction of memory deficits following surgical resection. However, there are two more investigational uses of the IAP that are worth discussing since they are employed by some epilepsy centers. The IAP has also been used to provide confirmatory evidence of lateralization of seizure focus as well as prediction of outcome in terms of seizure relief following surgery (Simkins-Bullock, 2000). The literature on these two uses will briefly be explored as they both relate to this study.

There is research supporting the use of the IAP to provide accurate confirmatory evidence of seizure lateralization. Perrine and colleagues (1993) reported on 57 individuals with temporal lobe epilepsy who were administered eight memory items. Verbal and non-verbal stimuli were used and left versus right memory performances were compared. Of the 57 individuals, 83% were correctly classified with poorer performance in the hemisphere ipsilateral to the seizure focus. In another study, Perrine and colleagues (1995) examined the relationship between seizure laterality and IAP Difference Scores in 70 individuals with temporal lobe seizure onset, left-hemisphere language, and no evidence of structural lesions. An IAP Difference score was defined as a memory score following right hemisphere injection minus the score following left hemisphere injection. This scoring system was able to classify 50 out of the 70 individuals into lateralized groups. Of the 50 individuals, 100% of the right seizure focus individuals and 96.4% of the left seizure focus individuals were correctly classified in terms of their seizure origin. In another study, Rausch, Babb, Engel and Crandall (1989) showed that impaired performance on the IAP memory test following injection to the hemisphere contralateral to the seizure focus strongly correlated with the hemispheric side of seizure focus in 19 of 30 individuals with intractable complex partial seizures. Wyllie and colleagues (1991) found similar results in 37 temporal lobectomy individuals with intractable partial epilepsy. Difference scores have also been used to predict lateralization in pediatric epilepsy surgery candidates with temporal lobe epilepsy. Lee and colleagues (2002) found differences in mean memory scores between left and right injections accurately predicted seizure-onset laterality to a statistically significant degree among the 87 children in the study who had unilateral TLE. In summary, there is evidence that the IAP can be helpful in predicting seizure lateralization and may be a useful adjunct to the neurodiagnostic evaluation.

There is also research to suggest the IAP is useful in predicting seizure outcome. One study reported on IAP memory asymmetries in 44 individuals who were seizure-free and 11 individuals who continued to have seizures one year following surgery. Asymmetries were defined as the disparity between memory scores during the left versus right hemisphere injection. When the authors grouped the individuals according to their ipsilateral-contralateral IAP asymmetry scores, significantly more individuals with large IAP memory disparities were seizure-free at follow up compared with individuals with small disparities. Out of the 44 patients, 36 had asymmetry scores of at least three points. Of these 36 patients 32 (89%) were seizure-free one year post surgery, and only 12 of the 19 patients (63%) of the patients with asymmetry scores less than 3 points were seizure-free at follow up (Loring, Meador, Lee, Nichols, et al., 1994). Perrine and colleagues (1995) examined the efficacy of memory difference scores in predicting postoperative seizure control in 70 individuals who underwent epilepsy surgery and had preoperative IAP memory testing. Outcome data for all 70 individuals were collected one year following IAP memory testing. The results indicated that IAP memory disparities were predictive of postoperative seizure control, with greater IAP memory asymmetries associated with better seizure control following surgery. Although there is evidence to suggest that the IAP is useful in predicting seizure outcome, additional outcome research is needed to establish the test as a predictor of seizure relief following surgical resection.

Chapter Four: The Purpose of This Study

Important treatment decisions are often influenced by IAP memory testing results, including the nature and extent of temporal lobe resection and, in some cases, whether surgery is performed at all. In the 1992 Palm Desert Survey, all 71 epilepsy centers reported they use the IAP to determine surgical parameters and approach (Rausch et al., 1993). Current IAP memory procedures vary widely by evaluating different types of memory (recall versus recognition) at different points in time (during anesthetization or post-drug effects), and with different methods of interpreting the resultant data. Consequently, different behavioral procedures might yield significantly different results and a patient may theoretically "pass" an IAP at one center and "fail" at another center. Recommendations from the 1992 Palm Desert Survey included the need to determine the validity of individual procedures and the need for a more standardized test among centers (Rausch et al., 1993). While standardization of procedures across centers is unlikely, there is a need for more centers to publish their IAP studies describing in detail their protocols, success rates for predicting postsurgical memory deficits, and other relevant validation data. If one center demonstrates validation data for their protocol, then other centers can adopt that protocol and greater uniformity in procedures across centers will hopefully result.

The primary purpose of this investigation was to validate an IAP memory assessment procedure designed using a statistical approach to help increase predictive value and reduce the probability of false positive and false negative errors. Reliability and validity data for this IAP memory assessment are presented, as well as correlations between IAP memory results and standardized neuropsychological assessment measures.

Hypotheses

Overall Goal: To investigate the psychometric properties of the IAP memory measure used at Parkland Health and Hospital System at UTSWMC in Dallas in an effort to demonstrate the reliability and validity of the test.

Specific Aim Number One: To investigate the overall reliability of three Forms of the IAP memory measure.

Hypothesis 1- Each form of the IAP memory test will demonstrate adequate internal consistency ($r \geq .70$) based on Cronbach's alpha.

Specific Aim Number Two: To investigate the psychometric properties of the IAP memory measure Forms I, II, and III through an examination of item-total correlations, item difficulty, item distractor analysis and item discrimination.

Hypothesis 2- All items on Forms I and II will have moderate difficulty (values between .20 and .80). In terms of difficulty, Form III will have the highest values, followed by Form I, and then Form II.

Hypothesis 3- Items on Forms I and II will demonstrate adequate discrimination ($D \geq .30$) between individuals who correctly recall $> 50\%$ of items correct and those who correctly recall $< 50\%$ of the items correct.

Hypothesis 4- All items on Forms I and II will have moderate item-total correlations ($r \geq .30$).

Hypothesis 5- For all items on Forms I and II, none of the seven distractors will be selected more than expected. To calculate the number of subjects expected to select each distractor, the number of subjects who answered the item wrong is divided by the number of distractors. Item distractor analysis will be conducted to investigate patterns of response on foils.

Specific Aim Number Three: To investigate the convergent validity of the IAP by investigating its relationship with standard neuropsychological tests of memory [i.e., the California Verbal Learning Test (CVLT), Rey-Osterrieth Complex Figure Test (Rey-O), and Wechsler Memory Scale-III (WMS-III)].

Hypothesis 6- The IAP memory scores on Forms I and II will significantly correlate ($r \geq .70$) with CVLT delayed recall, WMS-III total scores for Logical Memory II and Visual Reproduction II, and Rey-O delayed recall total score.

Hypothesis: 6a:

The IAP score for word items on Forms I and II will be more highly correlated with CVLT delayed recall and WMS-III Logical Memory II total score than Visual Reproduction II total score and Rey-O delayed recall total score. The IAP total score for picture items on Forms I and II will be more highly correlated with Visual Reproduction II total score and Rey-O delayed recall total score than CVLT delayed recall and WMS-III Logical Memory II total score.

Specific Aim Number Four: To investigate the divergent validity of the IAP by investigating its relationship with standard neuropsychological tests.

Hypothesis 7- The IAP total scores on Forms I and II will demonstrate low correlations ($r < .40$) with Trail Making Test parts A & B and Wisconsin Card Sorting Test number of perseverations.

Specific Aim Number Five: To explore the strength of the relationship between IAP performance and seizure lateralization.

Hypothesis 8 –Discrepancy Scores between left and right IAP injection performances on Forms I and II will predict seizure lateralization (based on clinical diagnosis).

Exploratory Hypotheses

Specific Aim Number Six: Examine the IAP memory measure's usefulness in predicting seizure outcome.

Hypothesis 9-Discrepancy Scores between left and right IAP injection performances on Forms I and II will be examined to see if there is relationship between higher discrepancies (in the expected direction) and better seizure outcome.

Chapter Five: Method

I. Subjects

Subjects included 90 consecutive individuals who underwent comprehensive neurodiagnostic assessment at Parkland Health and Hospital System between 1998 and 2005. Classification of individuals (i.e., left or right TLE) was determined by an epileptologist after interpretation of findings from a comprehensive neurodiagnostic evaluation including EEG, structural and functional neuroimaging techniques, IAP results, and neuropsychological assessment. Specific inclusion and exclusion criteria are listed below.

Inclusion Criteria:

1. Age 18 years of age or older.
2. Evaluated at the Epilepsy Monitoring Unit at Parkland Hospital.
3. Primarily English-speaking and evaluation with the IAP and neuropsychological testing in English.
4. Had a comprehensive neurodiagnostic evaluation.
5. Completed full 18 item IAP testing (i.e., baseline, left, and right injections).

<u>Exclusion Criteria:</u>

1. The IAP was considered invalid in the opinion of the attending epileptologist.
2. Individuals diagnosed with a mass lesion.

<u>IAP Memory Measure</u>

The IAP memory measure was developed at the University of Texas Southwestern Medical Center in 1998. The memory measure consists of 18 items composed of written words and pictorial items presented following injection of sodium amobarbital. The memory measure originally consisted of 16 items but two additional items were added to each form including a picture (i.e., The Taj Mahal) and a nursery rhyme (i.e., Roses are Red, Violets are Blue). Subjects are administered each item in the ipsilateral visual field during the period of hemiplegia and delta slowing on the EEG within 3 minutes of injection. Subjects are instructed to look at each picture and read each word aloud. The items alternate between written words (e.g., MINE, EXCEPT, INDEED) and pictorial items, which include both abstract shapes and everyday items (e.g., Cow, 4 curved arrows, safety goggles). After recovery of the hemiparesis and return of the EEG to baseline (approximately 10 minutes), recognition memory testing is conducted by having the patient select the target item presented earlier from eight items, seven of which are foils. All items are presented on an 8 ½ by 11 inch white paper. Recognition memory is tested with each correct item accompanied by 7 foils also presented on an 8 ½ by 11 inch sheet of white paper. The location of the correct response is varied across the 7 foils.

The IAP has three different forms with similar items. Form I is used following the first injection, Form II following the second injection, and Form III is used to obtain a baseline score. As a baseline, Form III was designed for all subjects to perform well. The number of items and foils were based on a binomial probability formula used to calculate the optimal number of foils for each tests with 16 to 18 items so that there was a low probability of obtaining a "passing score" by random guessing. Assuming a binomial distribution of guesses, individuals with no memory in the hemisphere being tested would have $<.02$ probability of obtaining a score of $>33\%$ correct by guessing on the test. The test is designed with sufficient difficulty to prevent both floor and ceiling effects so that individuals with moderate memory problems would have difficulty in selecting the correct item. During the design of the test, Form III was administered to 10 cognitively normal volunteers who all achieved a 100% recall score. The IAP was administered and scored by an attending epileptologist. See Appendix A for a complete copy of all three forms of the memory measure.

Neuropsychological Measures

The neuropsychological measures were chosen from a standard clinical epilepsy protocol used in the Neuropsychological Service at the University of Texas Southwestern Medical Center. With the exception of the IAP, all measures were administered and scored by experienced psychometricians or psychology interns who had no knowledge of the current study aims or hypotheses. The neuropsychological measures utilized in this study are described in Appendix B.

Procedure

Approval to conduct this study was obtained from the Institutional Review Board of the University of Texas Southwestern Medical Center at Dallas (UTSWMC) and written consent for clinical evaluation was obtained for all subjects. The subjects in this study underwent IAP testing at UTSWMC from May 1998 to January 2005 as part of a comprehensive presurgical evaluation. All individuals had prolonged surface ictal and interictal audiovisual and EEG recording. Interictal and ictal EEGs were reviewed by board-certified electrophysiologists. All individuals also underwent structural or functional imaging of the brain, using a technique sensitive for detecting sclerosis of the mesial temporal structures. Neuropsychological evaluation was conducted to help characterize neurobehavioral aspects of the seizure disorder. Administration of the complete test battery typically required four to six hours, and was usually completed in one day. While only the neuropsychological measures described in Appendix B were included in the statistical analyses, the entire battery was not limited to those measures. The neuropsychological battery generally included the following measures: Weschler Adult Intelligence Scale-Third Edition (WAIS-III), Wisconsin Card Sorting Test, Trail Making Test Part A and B, Rey-Osterrieith Complex Figure Test, California Verbal Learning Test (CVLT), and the Weschler Memory Scale-Third Edition (WMS-III) selected subtests, (Logical Memory and Visual Reproduction subtests), Digit Vigilance Test, Boston Naming Test, Oral Fluency Tests (FAS/Animals), finger tapping, and the Minnesota Multiphasic Personality Inventory-2. The order of the administration of tests was varied to meet the needs of each subject, with care taken to elicit maximum performance from all patients.

The IAP procedure involves angiographic visualization of the carotid circulation bilaterally to exclude anomalous vascular anatomy. Following angiographic visualization, 120 mg amobarbital, at a concentration of 20 mg/ml, is injected by hand over approximately 10 seconds into the internal carotid artery. In each case, the presumptive epileptogenic hemisphere was injected first (ipsilateral injection), followed 45 minutes later by presumptive nonepileptogenic hemisphere (contralateral injection). Individuals were examined clinically by an attending epileptologist, with continuous EEG monitoring, and the duration of hemiplegia, hemiparesis, hemianopia, aphasia, and EEG slowing was recorded.

The presence of obtundation and inattention was also noted. Tests were considered invalid if in the opinion of the attending epileptologist, obtundation or inattention was sufficient to interfere with registration of items during the presentation phase.

Subjects were then administered the 18 item test and items were presented in the ipsilateral visual field during the period of hemiplegia and delta slowing on the EEG. Items were presented within 3 minutes of injection. After recovery of the hemiparesis and return of the EEG baseline (approximately 10 minutes), recognition was tested by presenting the recognition stimuli involving each item accompanied by seven similar foils, and individuals were forced to choose one of the eight items. The test was scored as the fraction of items recalled correctly. Form II was used after the second injection, contralateral to the presumptive seizure focus. Form III was presented prior to the first injection to obtain a baseline score.

Statistical Procedures

The first aim of the study involved investigating the reliability of all three forms of the IAP memory test. Reliability refers to the consistency or dependability of measurements. The measure of reliability included in this study was internal consistency.

Hypothesis 1 stated that each form of the IAP will demonstrate adequate internal consistency ($r \geq .70$). Cronbach's coefficient alpha was used to measure the internal consistency of each form of the IAP. Cronbach's coefficient alpha is a numerical coefficient of reliability which ranges in value from 0 to 1. Alpha levels of .70 or higher are associated with increased test reliability as it indicates greater homogeneity and consistency of items (Nunnaly, 1978). Cronbach's alpha will be calculated for each item individually to assess the individual reliability of all 54 items.

Specific Aim Number Two involved investigating the psychometric properties of the IAP memory measure Forms I, II, and III through an examination of item-total correlations, item difficulty, item distractor analysis, and item discrimination.

Hypothesis 2 stated that all items on Forms I and II would have moderate difficulties (values between .20 and .80). In terms of difficulty, Form III would have the highest values, followed by Form I, and then Form II. Item difficulty was calculated for all 18 items on the two forms of the IAP memory measure. Item difficulty refers to the proportion or percentage of test takers who answer an item correctly. The higher this percentage is, the easier the item is (Crocker & Algina, 1986).

Hypothesis 3 stated that items on Forms I and II would demonstrate adequate discrimination ($D \geq .30$) between individuals who correctly recalled > 50% of items and those who correctly recalled < 50% of the items. An item discrimination index was calculated for all 18 items on the two forms. The item discrimination index, or D, is a measure of the effectiveness of an item in discriminating between high and low scorers on the test as a whole. The higher the value of D is, the more effective the item becomes. Items yielding positive discrimination index values of .30 or above, are considered good discriminators (Ebel, 1965).

Hypothesis 4 stated that all items on Forms I and II will have adequate item-total correlations. Item-total correlations were calculated for all 18 items on the two forms. An Item-total correlation is a correlation of each test item with the total score. Item-total correlations use a point-biserial correlation and range in values from +1 to -1. A positive biserial above .30 indicates the item is adequate (Crocker & Algina, 1986).

Hypothesis 5 stated that for all items on Forms I and II, none of the seven distractors would be selected more than expected. To calculate the number of subjects expected to select each distractor, the number of subjects who answered the item wrong was divided by the number of distractors. Item distractor analysis was conducted to investigate patterns of response on foils. Distractor analysis is a method of determining if foils are functioning as intended. Distractor analysis provides the frequency each foil is picked, and indicates the average number of subjects expected to select each distractor. The method also involves counting the number of times each foil is selected as the right answer by examinees in the upper group, middle group, and lower group. The lower group consisted of subjects who had less than 33% of the items correct, the middle group included subjects who had 34%-66% of the items correct, and the upper group consisted of subjects who had more than 66% of the items correct. Distractor analysis provides the percentage correct for all three groups, the percentage of distractors picked by all three groups, and the percentage from all three groups who picked the most frequently selected distractors.

Specific aims number three and four and hypotheses 6 and 7, involved investigating the validity of the IAP. Validity refers to the extent to which a test measures what it is supposed to measure. The two types of validity explored in this study were convergent and divergent validity. Convergent validity refers to the relationship between the IAP test scores and other related measures (in this case standard neuropsychological tests of memory). Divergent validity refers to the relationship between IAP test scores and unrelated measures. A Pearson correlation coefficient was utilized for all hypotheses and sub-hypotheses under specific aims three and four in order to measure the strength of relationships between tests. To further explore the convergent and divergent validity of the IAP memory measure, subjects injected on the right side first (left side second) were compared to subjects injected on the left side first (right side second) on the neuropsychological memory measures. Similarly, subjects with LTLE were compared to subjects with RTLE. This division of subjects into these subgroups allowed for easier comparison of left and right hemisphere memory performance on the IAP memory measure with neuropsychological testing.

Specific aim number five involved exploring the strength of the relationship between IAP and seizure lateralization.

Hypothesis 8 stated that Discrepancy Scores (DS) between left and right IAP injection performances on Forms I and II will predict seizure lateralization (based on clinical diagnosis). A discrepancy score or asymmetry score has been used in several studies involving IAP and accurate lateralization of seizure focus (Loring et al.,1994; Loring et al.,1995; Perrine et al., 1995). The discrepancy score (DS) used in this study was based on one described by Lee and colleagues (2002). As per Lee and Colleagues (2002) asymmetries were analyzed by using a difference score calculated by subtracting the percentage of items correctly recognized after the right injection from the percentage correctly recognized after the left injection (left memory minus right memory). DS can range from positive 100 to negative 100. Difference scores ≥ 20% suggest significant asymmetry with negative scores indicating better performance in the right hemisphere and positive scores indicating better performance in the left hemisphere. DS less than 20% were considered nonlateralized or indeterminate. These analyses were conducted with and without a correction factor. To compensate for lower performance as a result of aphasia, paraphasia, or mutism during left hemisphere injections. Perrine and colleagues (1995) applied a 17% correction factor to DSs in their investigation of the efficacy of DSs in predicting laterality. For this study, a 10% correction factor was applied to the DSs for all left injections. A 10% correction factor was selected as it provided the highest classification accuracy when compared to other correction factors (e.g., 5%, 15%, 20%). Logistic regression was used to explore the relationship between consensus diagnosis and the DS. Logistic regression is a predictive analysis of a dichotomous dependent variable. In this case, X is the DS finding (i.e., left or right seizure lateralization) and Y is the clinical diagnosis.

Exploratory Hypotheses

Specific aim number six was an exploratory analysis of the IAP memory measure's usefulness in predicting seizure outcome.

Hypothesis 9 involved investigating the use of Discrepancy Scores between left and right IAP injection performances on Forms I and II to see if there is relationship between higher discrepancies (in the expected direction) and better seizure outcome. DS were calculated as in the previous hypothesis with 20% used as the criterion for asymmetry significance. Outcome classification was based on Engel's four class postopertive classification. Class I refers to complete freedom from disabling seizures, Class II refers to rare disabling seizures, Class III indicates worthwhile improvement or prolonged seizure-free intervals, and Class IV refers to no worthwhile improvement. Outcome would be measured by reported seizure status 6 months post surgery. Class I was considered seizure free and Class II-IV were considered non-seizure free. (Engel & Shewmon, 1993). Outcome was examined qualitatively to see how discrepancy scores related to good vs. bad outcome.

Chapter Six: Results

Table 1 presents selected demographic characteristics of the sample. The 90 subjects who met inclusion criteria for the study included 39 males (43.3%) and 51 females (56.7%). The mean age was 35.6 years, and the mean years of education were 12.5 years. Excluded subjects included 43 subjects who were administered an older (16-item) version of the IAP memory measure and two subjects under the age of eighteen. Seventy-two subjects completed neuropsychological testing as part of their evaluation. The mean duration of seizures was 18 years and the mean estimated Full Scale IQ (FSIQ) was 87. Fifty-two subjects (57.8%) were Caucasian, 22 (24.4%) were Hispanic, 11 (12.2%) were African-American, two were Middle Eastern, and one was Asian. Seventy-three (81%) subjects were right-handed, and sixteen (19%) were left-handed. Seventy-nine (88%) subjects were left hemisphere language dominant, four (4%) were right hemisphere language dominant, and seven (8%) had bilateral language representation.

----Insert Table 1 here----

Seizure focus was determined by an epileptologist after interpretation of findings from a comprehensive neurodiagnostic evaluation including EEG, structural and functional neuroimaging techniques, IAP results, and neuropsychological assessment. Seizures were localized in the left hemisphere (n=42) more often than the right (n=22). Five subjects had bilateral seizure onset, and seizure lateralization was indeterminate after EMU evaluations in 21 subjects. Chi-square analyzes indicated no significant differences between sex and ethnicity among subjects with left versus right seizure lateralization. Fisher's Exact Test demonstrated no significant differences between handedness and language dominance in terms of subjects with right TLE (RTLE) versus Left TLE (LTLE). Independent-samples t tests indicated no significant difference in age, education, FSIQ or epilepsy duration, among subjects with left versus right seizure lateralization (Tables 2 and 3). There was also no significant difference in age, sex, ethnicity, handedness, education, FSIQ, or seizure duration between subjects injected on the right side compared to subjects injected on the left side first. There was no significant difference between subjects with an IQ < 85 (n = 33) and subjects with an IQ > 85 (n = 21) on Forms I, II, or III.

----Insert Tables 2 and 3 here----

Eighty-four subjects were diagnosed with TLE, 28 of whom were confirmed as having mesial temporal lobe epilepsy (MTLE). Six subjects had a nondiagnostic seizure evaluation. Table 4 presents the respective diagnoses as well as the diagnoses in terms of seizure lateralization.

----Insert Table 4 here----

Investigation of differences in performance across the IAP memory forms in relation to demographic variables showed that for Form I, Caucasians performed better than other ethnic groups, t (86) = 2.03, $p < .05$, and left handed subjects performed better than right handed subjects t (87) = -2.149, $p < .03$. Subjects with right language dominance performed better than subjects with left language dominance t (81) = 2.335, $p < .03$, and patients with a right seizure focus performed better than those with left seizure focus t (67) = 2.678, $p < .01$. There was a significant negative correlation between age and performance on Form I (r = -.33, $p < .01$), but there was no correlation between Form I and years of education, FSIQ, or disease duration.

On Form II, subjects diagnosed with a left seizure focus performed better than those with a right seizure focus t (67) = -4.511, $p < .01$. Performance on Form II significantly correlated with disease duration (r = -.34, $p < .01$), and FSIQ (r = .31, $p < .01$), but not age or years of education. There were no significant differences in performance across other demographic characteristics on Form II or Form III. Nor were there significant differences across demographic characteristics on Form III. Form III correlated with FSIQ ($r = .29$, $p < .05$) but not age, years of education, or disease duration. Table 5 presents the means and standard deviations for the raw scores for Forms I and II across sample characteristics.

----Insert Table 5 here----

Research Hypotheses

The first aim of the study involved investigating the reliability of all three forms of the IAP memory test. Hypothesis 1 stated that each form of the IAP tests would demonstrate adequate internal consistency ($\alpha \geq .70$). Cronbach's Alpha was utilized as a measure of internal consistency and had a value of $\alpha = .85$ for Form I, $\alpha = .83$ for Form II, and $\alpha = .69$ for Form III. The majority of items on the three forms contributed to test reliability as alpha values decreased when most items were deleted (see table 6). Thus, hypothesis 1 was supported for Forms I and II, and Form III was only slightly below .7.

----Insert Table 6 here----

Specific Aim Number two involved investigating the psychometric properties of IAP Forms I, II, and III through an examination of item-total correlations, item difficulty, item distractor analysis, and item discrimination.

Hypothesis 2 stated that all items on Forms I and II would have moderate difficulty (values between .20 and .80). Higher difficulty values are associated with a higher percentage of subjects answering the item correctly. In terms of difficulty, Form III was predicted to have the highest values, followed by Form I, and then Form II. Form III was expected to have the highest item difficulty values since it is a baseline measure and subjects are not injected with amobarbital during its administration. As predicted, the average item difficulty was .85 for Form III, .57 for Form I, and .41 for Form II. Form II was harder for all subjects regardless of injection side. For subjects injected on the left side first, the average item difficulty for Form I was .50 and .47 for Form II (second injection). For subjects injected on the right side first, the average difficulty for Form I was .68 and .31 for Form II or second injection (see table 7). For subjects injected on the left side first, there were six items on Form II that had item difficulty values less than .20. For subjects injected on the right side first, there was only one item with an item difficulty less than .20.

Items 13, 15 and 17 on Form I were greater than .80 and item 5 on Form II was less than .20. Following left injection on Form I, word items had a lower average item difficulty (.37) than picture items (.63). Following right injection, word items also had a lower average item difficulty (.66) than picture items (.70) on Form I, though the values were more similar. On Form II, following right injection, picture items had a lower average item difficulty (.45) than word items (.49). Following left injection, word items had a lower average item difficulty (.24) than picture items (.35) on Form II. Tables 8-10 present a complete summary of item difficulty across Forms I-III by first injection side. Hypothesis 2 was not completely supported, as it predicted that all the items would have moderate difficulties (values between .20 and .80) on Forms I and II (32 out of 36 items). However, most items on Forms I and II had moderate item difficulty values (89%), which is in within acceptable limits for such a test.

----Insert Tables 7, 8, 9, and 10 here----

Hypothesis 3 stated that items on Forms I and II would demonstrate adequate discrimination ($D \leq .30$) between individuals who correctly recalled > 50% of items and individuals who recalled < 50% of items correctly. An item discrimination index was calculated for all 18 items on the three forms. The item discrimination index, or D, is a measure of the effectiveness of an item in discriminating between high and low scorers on the test as a whole. The higher the value of D is, the more effective the item. Items yielding positive discrimination index values of .30 or above are considered good discriminators

(Crocker & Algina, 1986). Table 11 presents a listing of the item discrimination index scores for each item on all three Forms. On Form I, items 11, 13, and 14 had indexes below .3. Form II had no items with a discrimination value below .3. On Form III, a total of 13 items had item discrimination values below .3, which is not surprising since this is a baseline measure and most subjects performed at a high level. Form III had the largest range of values (0 to .57) followed by Form I (.22 to .74) and Form II (.32 to .83). Consequently, hypothesis 3 was fully supported for Form II and only partially supported for Forms I and III.

----Insert Table 11 here----

Hypothesis 4 stated that all items on Forms I and II would have moderate item-total correlations ($r \geq .30$). Item-total correlations use a point-biserial correlation and range in values from +1 to -1. A positive biserial correlation above .30 indicates the item is adequately measuring the same overall construct as the rest of the test. Table 12 presents the item-total correlations for all three forms. All items were predicted to be greater than .3, but on Form I, item 14 had an item-total correlation below .30 ($r = .22$); while on Form II, items 2, 5, 8, and 11 had item-total correlations below .30. There were eight items on Form III with item-total correlations below .3. Thus, hypothesis 4 was mostly supported for Form I and partially supported for Form II.

----Insert Tables 12 here----

Hypothesis 5 stated that for all items on Forms I and II, none of the seven distractors would be selected more than expected. To calculate the number of subjects expected to select each distractor, the number of subjects who answered the item wrong was divided by the number of distractors. The analysis indicated that on Form I items 5 and 12 had foils chosen three times more than expected and on Form II items 1 and 15 also had foils selected three times more than expected. On Form I, item 11 had a distractor that was chosen more than the correct response but not more than expected based on the formula provided above. On Form II, item 15 had a distractor that was chosen more than the correct response but not more than expected. Consequently, hypothesis 5 was mostly, but not completely supported for Forms I and II. In terms of distractors not selected by any subjects, on Form I, items 1, 8, and 17 had a foil that was not selected by any subject. On Form II, items 2, 4, and 15 had a foil that was not selected by any subject. Form III had the largest amount of items with foils not selected by any subject and the largest percentages of correct item selection among the lower group when compared to Forms I and II.

The third aim of the study involved investigating the convergent validity of the IAP memory measure. Hypothesis 6 stated that the IAP memory scores on

Forms I and II would be significantly correlated ($r \geq .70$) with CVLT delayed recall, WMS-III total scores for LM II and VR II, and Rey-O delayed recall total score. Pearson correlations indicated a significant relationship between Form I total score and CVLT delayed recall ($r = .39$, $p < .001$) as well as LM II ($r = .30$, $p < .009$). A significant relationship was also found between Form II total score and Rey-O delayed recall ($r = .36$, $p < .001$) and VR II ($r = 0.31$, $p < .008$). As indicated in table 13, Form I did not correlate with VR II or the Rey-O and Form II did not correlate with LM II and CVLT. Therefore, hypothesis 6 was not completely supported when looking at the whole sample of TLE patients who had neuropsychological testing. Table 14 includes the means and standard deviation for all subjects on the IAP memory measure as well as neuropsychological testing.

----Insert Tables 13 and 14 here----

The sub-hypotheses stated that total score for the word items on Forms I and II would more highly correlate with CVLT delayed recall and WMS-III LM II than with VR II and Rey-O delayed recall scores. Conversely, it was predicted that total scores for picture items on Forms I and II would more highly correlate with VR II total score and Rey-O delayed recall, than with CVLT delayed recall and WMS-III LM II.

The word item total score on Form I significantly correlated with LM II (r = .28, $p < .01$) and CVLT delayed recall (r = .39, $p < .001$), while correlations were not significant for VR II or Rey-O delayed recall. Word items from Form II were not correlated with CVLT delayed recall, LM II, or VR II, but were significantly correlated with Rey-O delayed recall (r = .31, $p < .007$). The Picture item total score from Form I significantly correlated with VR II (r = .28, $p < .02$) and CVLT delayed recall (r = .26, $p < .04$). The Picture item total score from Form II correlated with VR II (r = .36, $p < .002$) and Rey-O delayed recall (r = .34, $p < .003$), as well as LM II (r = .25, $p < .031$), but not CVLT delayed recall (see table 13). Thus, the sub-hypotheses were not completely supported when looking at the entire sample of patients.

The fourth aim of the study was to explore the divergent validity of the IAP memory measure by investigating its relationship with standard neuropsychological tests not associated with memory functioning. Hypothesis 7 stated that IAP total scores on Forms I and II would demonstrate low correlations (r < .40) with Trail Making Test parts A & B and the Wisconsin Card Sorting Test (WCST) number of perseverative responses. Form I total score was not correlated with Trail Making Test part A (r = -.07, p =.55), part B (r = -.04, p =.72) or WCST perseverations (r = -.01, p =.99). Form II total score had a low but significant negative correlation with Trail Making Test part B (r = -.26, p = .03) but was not correlated with Trail Making Test part A (r = -.09, *p* = .49) or WCST perseverations (r = -.11, *p* = .39). Therefore, this hypothesis was supported.

To further explore the convergent and divergent validity of the IAP memory measure, subjects injected on the right side first (left side second) were compared to subjects injected on the left side first (right side second) on the neuropsychological memory measures. Similarly, subjects with LTLE were compared to subjects with RTLE. The division of subjects into these subgroups allowed for comparisons between left and right hemisphere memory performance on the IAP memory measure with neuropsychological testing. Statistical analyses for these subgroups excluded six subjects who were not diagnosed with TLE. The first group of subjects examined included TLE patients injected on the right side first. Consequently, these patients received Form I after the first injection on the right side and Form II after the second injection on the left side. Table 15 depicts the means and standard deviations for the IAP memory measure scores as well as the neuropsychological memory test scores for this subgroup of subjects. Table 16 includes correlations between Forms I and II and CVLT, LM II, VR II and Rey-O delayed recall scores.

----Insert Tables 15 and 16 here----

For TLE subjects injected on the right side first (n = 31), Form I total score correlated with the CVLT delayed recall (r = .49, $p < .008$) as well as LM II (r = .57, $p < .001$). Form II total score correlated with VR II (r = .42, $p < .02$), but not Rey-O delayed recall. In terms of divergent validity, Form I total score was not correlated with Trail Making Test parts A & B or WCST number of perseverations. Form II total score negatively correlated with Trail Making Test part A (r = -.40, $p = .03$), but not WCST perseverations.

The word item total score on Form I was correlated with LM II (r = .53, $p < .003$) and CVLT delayed recall (r = .56, $p < .002$). There were no significant correlations between word item total scores from Form I and VR II or Rey-O delayed recall. The word item total scores on Form II were not correlated with any of the verbal or nonverbal memory tests. The picture item total scores from Form I were significantly correlated with LM II (r = .44, $p < .02$) but were not associated with any of the other memory tests. The picture item total scores from Form II were significantly correlated with VR II (r = .49, $p < .007$) but were not related to any other memory test.

Convergent and divergent validity was further explored in 53 TLE patients injected on the left side first. These patients received Form I after left injection and Form II after right injection. Table 17 provides the means and standard deviations for the IAP memory measure scores as well as the neuropsychological memory test scores. Table 18 includes correlations between Form I and II and CVLT, LM II, VR II and Rey-O delayed recall scores.

----Insert Tables 17 and 18 here----

For these subjects, Form I (left injection) did not significantly correlate with any of the memory tests. However, Form II (right injection) significantly correlated with Rey-O delayed recall (r = .34, $p < .03$). Form I word item total scores did not correlate with any memory tests, but Form II word item scores correlated with Rey-O delayed recall (r = .37, $p < .02$). The picture item total scores from Forms I and II were not correlated with any of the memory measures. In terms of divergent validity, Forms I and II did not correlate with the Trail Making Test or WCST.

Convergent and divergent validity was examined in the same sample by delineating subjects based on diagnosis of right versus left TLE. For subjects with RTLE (n = 22), Form I correlated with LM II (r = .55, $p < .009$) and the CVLT delayed recall (r = .45, $p < .04$), and Form II correlated with VR II (r = .44, $p < .04$) and the CVLT delayed recall (r = .48, $p < .03$). As with Form I total scores, Form I word item total scores correlated with LM II (r = .52, $p < .02$) and the CVLT (r = .53, $p < .01$). Form II total word item scores did not correlate with any memory tests. The picture item total scores on Form I did not correlate with any memory tests; however, the picture item total scores on Form II correlated with the VR II (r = .50, $p < .02$) and the CVLT (r = .49, $p < .02$). In terms of divergent validity, Forms I and II did not correlate with Trail Making Test parts A & B or WCST perseverations. Table 19 provides the means and standard deviations for the IAP memory measure scores as well as the neuropsychological memory test scores. Table 20 includes correlations between Form I and II and neuropsychological variables.

----Insert Tables 19 and 20 here----

When correlations for patients with LTLE (n = 42) were examined, Form I (left injection) did not significantly correlate with any of the memory tests; however, Form II (right injection) correlated with LM II (r = .37, $p < .04$) and the Rey-O delayed recall (r = .35, $p < .05$). Form I word item total scores did not correlate with any memory tests, though Form II word item total scores correlated with Rey-O delayed recall (r = .39, $p < .02$), the CVLT (r = .36, $p < .05$) and LM II (r = .39, $p < .03$) . The picture item total scores did not have significant correlations with any memory test. In terms of divergent validity, Forms I and II did not correlate with Trail Making Test parts A & B or the WCST perseverations. Table 21 provides the means and standard deviations for the IAP memory measure scores as well as neuropsychological memory test scores. In summary, there was some convergent validity for both Forms in subjects injected on the right first. For subjects injected on the left first there was some convergent validity for Form II but not Form I. Table 22 includes correlations between Form I and II and neuropsychological variables.

----Insert Tables 21 and 22 here----

Specific Aim Number Five involved exploring the strength of the relationship between IAP memory scores and seizure lateralization. Hypothesis 8 stated that Discrepancy Scores (DS) between the left and right IAP injection (right memory minus left memory) performances on Forms I and II would predict seizure lateralization (based on diagnosis). Logistic regression was used to see if DS predicted seizure lateralization. Difference scores $\geq$ 20% suggest significant asymmetry with negative scores indicating better performance in the right hemisphere and positive scores indicating better performance in the left hemisphere. DS less than 20% were considered nonlateralized or indeterminate. A 10% correction factor was added to DS scores following left hemisphere injections to compensate for lower performances associated with aphasia. These analyses were conducted with and without a correction factor and the DS were entered into the analyses as a continuous variable.

There were 42 subjects diagnosed with LTLE and 22 subjects diagnosed with RTLE. For subjects with RTLE, the mean DS was 45.45, and for subjects with LTLE, the mean DS was 1.11. For subjects with RTLE, all DS were positive and greater than 20% with or without a correction factor. For subjects with LTLE, 12 subjects (29%) had negative DS greater than 20%, 11 (28%) had falsely lateteralized DS greater than 20% and 19 had indeterminate scores. When applying the correction factor, 16 subjects (38%) had negative DS greater than 20%, 7 (17%) subjects had positive DS greater than 20%, and 19 (45%) subjects had indeterminate scores (Table 23). When entering the DS (without the correction factor) into the equation, the logistic regression was statistically significant and the model was able to correctly classify 83% of the subjects in terms of lateralization. The Wald statistic, which determines if any of the independent variables in the regression equation has a significant relationship with the dependent variable using a conventional significance level of $p \leq .001$, demonstrated that the DS score significantly contributed to the regression model χ^2 (1, N = 64) = 14.84, $p < .001$. However, the Hosmer and Lemeshow Goodness of Fit Test showed this regression equation's model had a limited fit with the observed data χ^2 (8, N = 64) = 14.32, $p < .08$. When the logistic regression for the DS was compared to logistic regression analyses when Form I and II total scores were both entered into the equation, the total scores had nearly the same percentage of correct classification (80%) and the Hosmer and Lemeshow

Goodness of Fit Test showed that this regression equation's model had a good fit with the observed data χ^2 (8, *N* = 64) = .121, p <.95. The Wald statistic was significant for both Form I, χ^2 (1, N = 61) = 10.60, p < .001 and Form II (1, N = 61) = 10.99, p < .001.

When a correction factor was employed to see if the DS became a better predictor of lateralization, the logistic regression was statistically significant and the model was able to correctly classify 92% of subjects in terms of lateralization. The Wald statistic demonstrated that the DS score significantly contributed to the regression model χ^2 (1, N = 64) = 10.80, p < .001. When using the correction factor, the Hosmer Lemeshow Goodness of Fit test showed this regression equation's model had a good fit of the observed data χ^2 (8, *N* = 64) = 5.17, *p* = .74. Consequently, the DS (when using the correction factor) was a better predictor of lateralization than Form I and II total scores by themselves. Therefore, Hypothesis 8 was supported by the logistic regression, as the corrected DS was able to classify 92% of subjects. Table 23 includes the mean DS for the patients with right and left TLE.

----Insert Tables 23 here----

Exploratory Hypotheses

Specific Aim Number Six is an exploratory analysis of the IAP memory measure's usefulness in predicting seizure outcome. Hypothesis 9 involved investigating the use of Discrepancy Scores between left and right IAP injection performances on Forms I and II to see if there was relationship between higher discrepancies (in the expected direction) and better seizure outcome. DS were calculated as in the previous hypothesis with 20% used as the criterion for asymmetry significance. Among the TLE patients, there were 40 subjects who had outcome data and 44 subjects with missing data. A total of 26 subjects (68%) had complete freedom from disabling seizures (Class I) and 14 subjects (35%) had at least one post-operative seizure. For seizure free subjects, 21 (81%) had DS greater than 20%. For non seizure-free subjects, 8 (57%) had DS greater than 20% and 6 had DS less than 20%. Seizure free subjects had a significantly larger mean discrepancy score [42.74 (t (38) = 2.167, p < .05)] than non seizure-free subjects (27.38). In addition, when comparing performance means on Form I there was a significant difference between seizure free versus non-seizure free patients on Form I with TLE t (38) = 2.264, p < .05.

Chapter Seven: Discussion

Over 400,000 individuals suffer from temporal lobe epilepsy and for many, seizures are refractory even with pharmacological intervention. Surgical resection is the most promising treatment for intractable TLE, as 60%-80% of patients are rendered seizure free following temporal lobectomy (Devinsky, 2004). The goal of surgical resection is to remove the seizure focus to substantially reduce or eliminate seizures. IAP is considered an essential part of the presurgical neurodiagnostic evaluation in most epilepsy centers throughout the country (Rausch, 1993). The IAP's original purpose was to establish language dominance, but in 1962, Brenda Milner and colleagues extended the use to include testing of memory functions. Today the IAP continues to be routinely used to predict memory deficits following surgery. However, many epilepsy centers also utilize the IAP to provide confirmatory evidence regarding seizure lateralization and prediction of seizure relief following surgery (Simkins-Bullock, 2000).

Even with the IAP's mainstream use, there is great variability in how the test is being performed across centers. The inconsistencies in procedure include amount of amobarbital injected, drug injection procedure, the criteria for determining whether the drug has taken effect and the definition of "passing" versus "failing" the test. In addition, there is wide variability in the memory measure of the IAP including the nature of the stimuli presented (i.e., real objects or line drawings), the timing of memory stimuli presentation, the number of stimuli presented, and the use of free recall versus recognition (Simkins-Bullock, 2000).

Despite the inconsistencies in procedures across centers, it is not uncommon for treatment decisions regarding surgical resection to be influenced by the results from IAP testing and in some cases, patients may be denied surgery if the IAP is suggestive of risk of amnesia following resection (Rausch et al., 1993). The high success rate of anterior temporal lobectomy (ATL) has led to its wide acceptance as a treatment of choice, although it is still underutilized. The number of patients undergoing ATL represents only a small percentage of the potential candidates for surgery (Engel & Shewmon, 1993). Determination of neuropsychological outcome continues to be a major factor in deciding whether to proceed with ATL. The majority of cognitive abilities remain stable following ATL (Chelune, Naugle, Luders, Sedlak, & Awad, 1993); and postsurgical amnestic disorders are rare, although significant decline in episodic memory is sometimes associated with ATL (Loring et al., 1994; Rausch et al., 1993).

The value of IAP in predicting post-surgical memory deficits has been determined in several studies (Chelune, et al., 1993; Kneebone et al., 1995). There are two theories regarding prediction of post-surgical memory deficits that are frequently considered. The functional reserve model asserts that postsurgical memory deficits depend on the integrity of the temporal lobe contralateral to the side of resection to support memory function following surgery. The hippocampal adequacy model states that postsurgical memory deficits are dependent on the functional capacity of the tissue to be resected (Chelune, 1995). In most epilepsy centers, the side of the presumed seizure focus is injected first in case of procedural complications that prevent second injection (Acharya & Dinner, 1997). The 1992 Palm Desert Survey indicated that 36% of epilepsy centers use evidence regarding intact contralateral memory to consider modifying surgical approach (Rausch et al., 1993). Therefore, the functional reserve model appears to be an important factor in determining postsurgical memory functioning.

The main purpose of this study was to validate a standardized IAP memory test and explore its value in predicting lateralization as well as treatment outcome. Four components of test validation were examined in the current study: 1) The overall reliability of the measure was examined as well as the reliability of the individual items. 2) A comprehensive examination of the 54 items and 378 foils that comprise the memory measure. This item analysis included item difficulty, item discrimination, item-total correlation, and distractor analysis. 3) Correlations with other neuropsychological memory measures were conducted to demonstrate convergent and divergent validity. 4) Examination of the utility of discrepancy scores in predicting seizure lateralization and surgical outcome. The hypotheses in this study were based on criteria which define characteristics of a "good test." In the ensuing sections, the findings and limitations of the study will be discussed and implications for further research will be addressed.

In terms of demographic variables, the group of subjects injected on the left side first was comparable to the group injected on the right side first. In terms of race, Caucasians performed significantly better than other ethnicities on Form I. In addition, left handed subjects performed better than right handed subjects on Form I. In terms of overall performance, subjects diagnosed with RTLE performed better than subjects with LTLE on Form I. This is not surprising, as LTLE subjects were typically injected first on the left side first, and left injections have been associated with worse IAP performance (Grote et al., 1999). Subjects with LTLE performed better than subjects with RTLE on Form II. This is not unexpected, as their second injection was on the right side, and left injections often negatively impact performance. For subjects with RTLE, the total mean score was significantly worse (t (21) = 13.94, p < .05) for Form II (4.32) compared to Form I (12.5). Such a large discrepancy may be related to the impact during the second injection of testing the affected hemisphere in addition to aphasia during memory testing. FSIQ scores did not correlate with Form I, but had a small correlation (r = .29 to .31) with Forms II and III. There was no significant difference in total scores between subjects with a FSIQ greater than 85 and subjects with a FSIQ less than 85 on any of the Forms from the IAP memory measure. Thus, FSIQ scores did not appear to have a significant impact on results.

<u>Reliability</u>

Reliability refers to the consistency of measurement. One aspect of reliability is internal consistency, which refers to the degree individual items measure a similar construct (Anastasi, 1988). The first aim of the study involved investigating the reliability of all three Forms of the IAP. Hypothesis 1 stated that each form of the IAP would demonstrate adequate internal consistency ($r \geq .70$) based on Cronbach's alpha. As predicted, both Forms I and II had high reliability with alpha values of .85 for Form I, and .83 for Form II. Nunnaly (1978) has indicated .7 to be an acceptable reliability coefficient, but lower thresholds are sometimes used in the literature. As indicated in Table 6, the majority of items on Forms I and II contributed to the reliability of the test with the exception of Item 2 on Form II ("These"). For this item, there was a small increase (.01) in reliability when the item was removed. This item also had a low item-total correlation ($r = .17$) and may be testing a different construct from the overall test.

Since Form III is without amobarbital injection, the majority of subjects obtained near perfect scores; therefore, there was less variance among scores. As indicated in Table 6, Form III had several items in which the reliability of the Form increases if those items are deleted. This is most likely a result of the limited variance of this baseline Form. Cronbach's alpha is related to the variance of the item scores as well as total test variance. If there is less variance in the items as well as less variance in the total test scores, alpha levels will be lower. Even with less variance, Form III maintains an alpha level (.69) close to the .7

standard for psychometric tests. Cronbach's alpha measures the uniformity, or homogeneity of items throughout the test. Therefore, based on the alpha levels across the three forms, the IAP memory procedure appears overall to have good reliability and internal consistency. Reliability estimates remained above .8 regardless of FSIQ score, with the exception of Form III. For Form III, Cronbach's alpha dropped to .6 in subjects with a FSIQ > 85. This is not surprising as these subjects answered most items correctly; limiting the variance of item scores and negatively impacting the reliability of the measure.

Item Analysis

Specific aim number two involved investigating the psychometric properties of Forms I, II, and III through an examination of item-total correlations, item difficulty, item distractor analysis, and item discrimination. Hypothesis 2 stated that all items on Forms I and II would have moderate difficulty (values between .20 and .80). It was also predicted that Form III would have the highest values, followed by Form I, and then Form II. A higher item difficulty value is associated with a greater percentage of subjects answering the item correctly. As predicted, Form III had the highest mean item difficulty values, followed by Form I, and then Form II. Form III was predicted to have the highest item difficulty values since it is a baseline measure in which there is no injection of sodium amobarbital (i.e., subjects are supposed to do well at baseline). Form II had the lowest item difficulty value regardless of injection side, most likely because Form II is usually testing the affected hemisphere.

In terms of overall item difficulty, several items on Forms I and II had difficulty values greater than .8 or less than .2. Items 13, 15 and 17 on Form I were greater than .80 and item 5 on Form II was less than .20. On Form I, Item 13 (picture of a gymnast on a balance beam) had an item difficulty value of .84 suggesting it was an easier item, as at least 84% of subjects picked the item regardless of side of injection. Items 15 (picture of a 4 star flag) and 17 (picture of the Taj Mahal) on Form I, also had item difficulty values greater than .8. Thus, these three items may be candidates for revision as their item difficulty values suggest that more than 80% of subjects answered the item correctly. For subjects injected on the right first, six items had item difficulty values greater than .8. Four of these items were pictures and two were word items. Since verbal memory is considered to be a function of the left cerebral hemisphere and visual memory a function of the right cerebral hemisphere (Kimura, 1963; Kolb & Whishaw, 1990) it is surprising that higher item difficulty values would be associated with picture items when testing left memory. Since five of these items (13, 15, 16, 17, and 18) occur later in the test, it is possible that the amobarbital's effect had dissipated to some degree. Most epilepsy centers present stimulus items within the first three minutes following amobarbital injection to avoid presentating items without the effect of the amobarbital. Loring and collegues (1994) demonstrated that stimulus items presented 30 to 45 seconds following injection were significantly better at predicting seizure laterality than items presented later during the IAP. However, in this study, higher item difficulty values were not consistently observed in later items. Thus, these five items may just be less difficult for subjects injected on the right side first.

For subjects injected in the left hemisphere first, the item difficulty values were all in the moderate range. The most difficult item was Item 11 on Form I (picture of China) which had an item difficulty value of .21. This particular item did not discriminate well (D = .28) between the upper and lower performing groups and it had a low item-total correlation. This item will be discussed further in a later section.

On Form II, Item 5 (picture of the state of Virginia) is a good candidate for revision as it is had an item difficulty value of .12 regardless of side of injection. This item was the most difficult item on the IAP memory measure, as only 32% of subjects in the upper performing group (top 33% of subjects) were able to get this item correct, which is in contrast to the upper performing group's performance on all other items on Form II (range of percentage of items correct = .32 to .97). For subjects injected on the right side second, item difficulty values were in the moderate range with the exception of item 5 (Form II). For subjects injected on the left side second, six items had item difficulty values below .20. One of those items was item 5 (picture of Virginia) which was previously discussed. The other five items were word items. It is not surprising that these items had low item difficulty values since the right affected hemisphere was being tested (i.e., performance negatively impacted by loss of language with left injection and testing affected hemisphere). Since the IAP is also used to assess language dominance, left injection usually results in global aphasia, paraphasias or muteness for a brief period provided the left hemisphere is language dominant (Rausch et al., 1993). Thus, it is not surprising that word items on Forms I and II had lower average item difficulty values following left injection than picture items.

Form III had an average item difficulty value of .85 and contained six items with item difficulty values above .95, suggesting that almost all subjects obtained near perfect scores on these items. The most difficult item on Form III ("after") had an item difficulty of .49, while Item 7 (picture of a grape cluster) was correctly selected by all subjects. Overall, the majority of items on Forms I and II had moderate difficulty values, supporting the second hypothesis and the use of the IAP memory measure.

Item discrimination is an evaluation of how individual items discriminate between the group of subjects who performed well on the test (the upper performing group) and the group of subjects who performed poorly on the test (the lower performing group). Hypothesis 3 stated that items on Forms I and II would demonstrate adequate discrimination ($D \geq .30$) between individuals who correctly recalled > 50% of items and those who correctly recalled < 50% of the items.

On Form I, items 11, 13, and 14 had item discrimination indexes below .3. Item 11 (picture of China) had a difficulty value of .24 indicating that it was a difficult item for subjects. This was most likely the reason it did not discriminate well between groups. In fact, more subjects selected the foil with a picture of Brazil over the correct answer. This item's low discrimination and difficulty value may be related to the similarity of the foils. As mentioned earlier, item 13 had an item difficulty value of .84, signifying most subjects correctly identified it. This is the reason this item has a low discrimination value. Item 14 on Form I ("between") had an item difficulty value in the moderate range (.41), but had a similar percentage of subjects in the lower performing group selecting that item as in the upper performing group. In addition, the low item-total correlation for item 14 on From I ($r = .22$) suggests it is tapping a different construct from the rest of the test. Therefore, items 11, 13, and 14 on Form I may be candidates for revision. Form II had no items with a discrimination value below .3 and Form III had eight items with discrimination values below .3. Given that subjects got most of the items correct on Form III, there was less of a range in performance on items and consequently, low discrimination values. Overall, 33 out of 36 items on Forms I and II had acceptable discrimination values, supporting the hypothesis and the use of the IAP memory measure.

The next hypothesis involved item-total correlations which assess the degree to which items are measuring the same construct as the overall test. Higher correlations are associated with greater test reliability and internal consistency. High item-total correlations also indicate that the items discriminate examinees in the same direction as does the whole test. Hypothesis 4 stated that all items on Forms I and II would have moderate item-total correlations ($r \geq .30$). On Form I, item 5 (picture of Washington) and item 14 ("restore") had item-total correlations below .3. On Form II, items 2 ("these"), 5 (picture of Virginia), 8 ("hanger"), and 11 (picture of Africa) had item-total correlations below .30. Items 5 and 11 on Forms I and II consist of geographical locations (Washington, China, Virginia and Africa, respectively). The foils for these items appear less distinct than other items which may be the reason they are harder for subjects to discriminate. They also appear more abstract or less identifiable than other foils. These items may be good candidates for revision since three of them are testing different constructs than the overall test, and item 11 on Form I had a low discrimination value. Form III had eight items with item-total correlations below .3. This is not unexpected as the items on Form III have limited variance which creates a restricted range. Correlations are smaller when there is a restricted range (Heiman, 1996).

Hypothesis 5 stated that for all items on Forms I and II, none of the seven distractors would be selected more than expected. To calculate the number of subjects expected to select each distractor, the number of subjects who answered the item wrong was divided by the number of distractors. This hypothesis was not entirely supported as there were a few foils selected more than expected. On Form I, item 5 (picture of Washington) and item 12 ("lantern") had one foil that was selected much more than expected. For item 5, the average number of subjects expected to choose each distractor was 7. However, the last foil (picture of New York) was selected 21 times and thus stands out or appears more distinct than other foils. For item 12, the average number of subjects expected to choose each distractor was 7. Yet, the foil "elevator" was selected 24 times. On Form II, item 1 (picture of safety goggles) and item 15 (picture of Betsy Ross Flag) had foils selected much more than expected. For item 1, the average number of subjects expected to select each distractor was 8. The first foil (picture of hammer) on item one was selected 30 times. This item was most likely confused with item 4 on Form I (the word "hammer"). For item 15, the average number of subjects expected to choose each distractor was 9. The first foil (picture of flag with 13 stars) on item 15 was selected 36 times. Distractor analysis also indicated items in which foils were not selected by any subject. Usually, if foils are not selected, they are easy for subjects to discount and may be too dissimilar from the correct item. On Form I, items 1 (picture of cow), 8 ("river"), and 17 (picture of Taj Mahal) had foils that were not selected by any subject and on Form II, items 2 ("these"), 4 ("apple"), and 15 (picture of Betsy Ross Flag) had foils that were not selected. For Form III, almost every item had a foil that was not selected. This is not surprising as most subjects selected the correct answers. Although these foils are selected much more or less than expected, most are not among the problematic items discussed earlier and are not negatively impacting the overall reliability of the test.

In summary, the IAP memory measure had good overall reliability for all three Forms (Cronbach's Alpha = .85 for Form I, .83 for Form II, and .69 for Form III). There were some items on Form I (5, 11, 13, 14, 15, and 17) and a few items on Form II (2, 5, 8, and 11) that may benefit from revision, but most items on both forms contributed to the reliability of the test, and overall, the majority of hypotheses with respect to specific aims one and two were supported.

Convergent and Divergent Validation

Determining if a test is measuring the construct it was designed to measure is an important aspect of test validation. Two components of construct validation are convergent and divergent validity. Convergent validity is represented by the positive correlation of the IAP memory measure with scores from other neuropsychological memory measures. Conversely, divergent validity is demonstrated by low correlations with tests that do not measure memory skills. Hypothesis 6 stated that the IAP memory scores on Forms I and II would be significantly correlated ($r \geq .70$) with CVLT delayed recall, WMS-III total scores for LM II and VR II, and Rey-O delayed recall total score. When looking at these hypotheses with all subjects, various factors complicated the interpretation of correlations. Subjects varied in terms of side of injection, side of seizure lateralization, side of language dominance, and diagnosis. To simplify interpretation, correlations were limited to subjects diagnosed with TLE with left or mixed language dominance (n = 86). In addition, subjects were investigated in terms of side of injection and side of seizure focus. The division of subjects into these subgroups allowed for greater understanding of left and right hemisphere performance on the IAP memory measure in relationship to the neuropsychological testing. Thus, the hypotheses from specific aims three and four were explored in terms of side of amobarbital injection (i.e., left injection first, right injection second) and side of seizure focus (i.e., LTLE or RTLE).

The sub-hypotheses stated that the total score for word items on each of Forms I and II would be more highly correlated with CVLT delayed recall and WMS-III LM II, than VR II and Rey-O delayed recall scores. Conversely, it was predicted that total scores for picture items on each of Forms I and II would be more highly correlated with VR II and Rey-O delayed total, than with CVLT delayed recall and WMS-III LM II. Since verbal memory is considered to be a function of the left cerebral hemisphere and visual memory a function of the right cerebral hemisphere (Kimura, 1963), it was assumed that the word items would more highly correlate with verbal tasks following right injection (left memory) and the picture items would more highly correlate visual tasks following left injection (right memory). However, this assumption was not consistently found.

When looking at the correlations for subjects injected on the right side first or with RTLE, the hypotheses are largely supported. When injecting the right side first and testing the left hemisphere, the word item total score from Form I correlated with LM II and CVLT delayed recall. Although correlations were lower than predicted, this finding provides some validation for the word items on Form I as they were related to the verbal memory tasks. The picture items on Form I correlated with LM II. This may indicate that the picture items on Form I are being verbally encoded. Following second left injection, picture items on Form II (testing right memory) correlated with VR II but not Rey-O. This provides some convergent validity for the picture items on Form II. In addition, in the RTLE group, the picture items correlated with the CVLT. This is another indication that some picture items may have been verbally encoded.

In both hypothesis 6 and the sub-hypotheses, it was predicted that the IAP memory test would be highly correlated with the neuropsychological memory measures ($r \geq .70$). However, the highest correlation was $r = .5$, and correlations ranged from $r = -.03$ to $r = .57$. Correlations around .5 indicate a moderate relationship between variables. One possible explanation for the lower correlations with the neuropsychological memory tests is that the IAP memory measure is a different task as compared to the neuropsychological memory tests.

Unlike the IAP memory measure, the neuropsychological memory tests assess verbal and visual memory in different subtests. The IAP memory measure also involves a recognition format. The correlated components of the neuropsychological tests were not recognition formats. In addition, the neuropsychological memory tests were administered without interference of amobarbital. In addition, correlations between the neuropsychological measures themselves are not above .7. The Psychological Corporation (1997) examined correlations between WMS-III, CVLT, WMS-R, and Rey-O in several clinical groups. The WMS-III Auditory Delayed Index is composed of LM II total score and Verbal Paired Associates II Recall total score. This index had a .5 correlation with the CVLT. The WMS-III Visual Delayed Index (composed of Faces II and Family Pictures II) had a low correlation (r =.22) with the Rey-O delayed total score. In other studies, correlations with visual memory tests and the Rey-O have ranged from .45 to .65 (Knight & Kaplan, 2003). Thus, only modest correlations between some of the neuropsychological memory measures themselves have been demonstrated. In addition, low correlations may have resulted from the limited range of performance on the IAP memory measure under some conditions. For example, raw IAP scores were particularly low for RTLE subjects during second injection (mean total score was 4.32). This low performance on the IAP memory measure may have impacted the correlations due to the restricted range of scores. Another reason the correlations were lower than expected may be related to the time interval between neuropsychological assessment and the IAP assessment. The mean interval was 7 months but ranged from 0 to 51 months. Thus, it is possible that this time interval between assessments may have impacted certain correlations. Overall, it appears that the prediction of correlations greater than .7 was an overestimate and somewhat unrealistic.

As mentioned earlier, the correlations between picture items on Forms I and II with verbal memory tests such as CVLT and LM II may indicate that the picture items were being verbally encoded. For subjects with RTLE, the picture items were examined individually. Item 15 shows a significant correlation with CVLT ($r = .83$, $p < .001$). When the picture item total score from Form II is correlated without item 15, there is no significant correlation with CVLT in patients with RTLE ($r = .26$, $p = .180$). Item 15 on Form II was not a problematic item for the IAP memory test. The item portrays a picture of the Betsy Ross flag and the high correlation provides evidence that the item may be verbally processed. In terms of neuropsychological assessment, the development of truly “nonverbal” stimuli has been challenging. In one study, Eadie and Shum (1995) explored the utility of Chinese characters to assess nonverbal memory. They found Chinese characters to be a better measure of visual memory than geometric figures. Thus, the creation of stimuli that truly measure nonverbal memory continues to be a challenge.

Convergent validity was further examined in subjects injected on the left first as well as subjects diagnosed with LTLE. When testing right memory, Form I had no significant correlations with any of the memory tests. Similarly, Form I word item total score and the picture item total score were unrelated to the neuropsychological memory tests. This finding was somewhat surprising. One possible explanation for this finding is that Form I is not a good measure of right hemisphere function. However, it is also possible that some aspects of visual memory may be processed in the left hemisphere as well. Several studies have demonstrated verbal deficits in patients with TLE but have failed to identify nonverbal memory decrements (Barr, 1997; Chelune et al., 1991; Delaney, et al., 1980). In addition, Larrabee and colleagues (1985) were unable to demonstrate the presence of a nonverbal memory factor when examining the factor structure of the WMS-R. In addition, some studies have alluded to the fact that the construct of "pure" right hemisphere nonverbal memory may not exist (Larrabee et al., 1985; Smith et al., 1992). In general, studies have supported verbal memory decline following left ATL. Consistent results have been found using verbal list tasks (Herman, Wyler, Bush, & Tabatabai, 1992; Mungas, Blunden, Benington, Stone, & Palma, 1990), and paragraph recall measures (Delaney, Rosen, Mattson, & Novelly, 1980). Nonverbal memory tasks have been much less successful at identifying impairments in patients with right TLE (Naugle et al., 1993). In fact, the WMS-R Visual Reproduction subtest has not consistently demonstrated differences in VR subtest scores between patients with left or right TLE (Chelune et al., 1991; Glowinski, 1973); Ivnik et al., 1987). One study found no relationship between the Visual Reproduction subtest of the WMS-R and right hippocampal sclerosis (Rausch and Babb, 1993). Thus, the overall difficulty in measuring right hemisphere function has been demonstrated in the literature and is likely to be an obstacle for most IAP memory procedures.

When testing left memory following second injection, Form II was unrelated to the verbal memory tests, but demonstrated a low, but significant, correlation with the Rey-O. In subjects with LTLE, The word item total score from Form II (testing left memory) correlated with LM II and CVLT in addition to Rey-O. Thus, there appears to be some convergent validation for the word items on Form II in subjects with LTLE; however, the correlations are much lower than expected.

The fact that the Rey-O correlated with the word items may indicate that aspects of the Rey-O stimuli can be verbalized. The Rey-O and other visual memory tasks have been criticized for their susceptibility to verbalization (Heilbronner, 1992). The Rey-O has failed to lateralize epilepsy in several studies (Knight & Kaplan, 2003). Barr and colleagues (1997) compared the Rey-O performance of 187 LTLE patients with 168 patients RTLE, and found no significant differences across seven epilepsy centers. Taylor (1969) attempted to create a more “complex” visual figure than the Rey-O, but found no differences in performance between patients with left and right TLE. However, he did find that the right TLE group recalled significantly fewer details after a 40 minute delay interval. Miller and colleagues (1993) found decreased recall of details of the complex figure in patients with hippocampal sclerosis, but found no differences between right and left temporal groups.

Specific aim four involved examining the divergent validity of the IAP memory measure. Divergent validity reflects the adequacy of a test by showing insignificant relationships to measures that are extraneous to the target construct (Knight & Kaplan, 2003). Therefore, in Hypothesis 7, it was predicted that the IAP total scores on Forms I and II would demonstrate low correlations ($r < .40$) with the Trail Making Test and the Wisconsin Card Sorting Test. In subjects injected on the right side first, Form I total score was not correlated with Trail Making Test or the WCST and supports the divergent validity of this Form. Form II total score negatively correlated with Trail Making Test part A ($r = -.40$, $p < .03$), but was not associated with Trails part B or WCST perseverations. This negative correlation suggests that better performance on Trails A was associated with higher scores on Form II. This finding does not support hypothesis 7; however, the correlation is only slightly greater than .4. In addition, Trails A is a measure of visual attention and the IAP memory measure involves visual attention. Thus, Trails A may not be the best measure of divergent validity. For subjects injected on the left side first, Forms I and II did not correlate with Trails

Making Test parts A & B or WCST perseverations supporting the construct validity of the IAP memory measure. There were also no significant correlations between Forms I and II and Trail Making Test A & B or WCST perseverations in the right and left TLE groups. Overall, divergent validity is provided by the low and mostly insignificant correlations with these neuropsychological tests.

In summary, word items on Forms I and II demonstrated some convergent validity, through correlations with the verbal memory tests. The correlations were well below expectation; however, the prediction of high correlations ($r > .7$) may have been unrealistic. The picture items on Form II demonstrated some convergent validity through correlations with VR II. Conversely, the picture items on Form I failed to correlate with the visual memory tests which may represent the difficulty in developing stimuli that assess “nonverbal” memory. In addition, this finding may be an indication that nonverbal memory involves more than just the right hemisphere. For example, some studies have demonstrated more anterior brain regions in localization of nonverbal forms of working memory (Courtney, Petit, Maisog, Ungerleider, & Haxby, 1998). In terms of divergent validity, the IAP memory measure was validated through low correlations with neuropsychological measures unrelated to memory. Overall, there appears to be evidence for the construct validity of the IAP memory measure.

Discussion of Specific Aims Five and Six

Some investigators have found difference scores between left and right injection performances on the IAP to be associated with accurate lateralization of seizure focus (Loring et al.,1994; Loring et al.,1995; Perrine et al., 1995). In this study, a DS (as described by Lee and colleagues, 2002) was calculated by subtracting left memory IAP scores from right memory IAP scores. For patients injected on the right first, the DS were calculated by subtracting the percentage of items correctly identified on Form I from the percentage of items correctly recognized on Form II (left memory minus right memory). Difference scores ≥ 20% were defined as a significant asymmetry, with negative scores indicating better performance in the right hemisphere and positive scores indicating better performance in the left hemisphere. DS less than 20% were considered nonlateralized or indeterminate. Aphasia, paraphasia, and mutism often occur following left injection and consequently performance on the IAP has been shown to be adversely impacted during left injection. Without use of a correction factor, the logistic regression was able to classify 82% of subjects, but the model did not have a good fit with the data. When applying a correction factor to those subjects with lateralization data, the logistic regression was able to classify 92% of subjects with right or left TLE, and the model had a much greater fit with the data. As indicated in the results section, all subjects with RTLE had DS greater than 20% that were in the correct direction. Among the subjects with LTLE, 16 (39%)

had DS greater than 20% in the correct direction and 7 DS were falsely lateralized. It is not surprising that DS falsely lateralized subjects with LTLE. Subjects with LTLE frequently demonstrate more wide spread memory deficits than subjects with RTLE. Perrine and colleagues (1995) examined the use of DS and with a 17% correction factor, and were able to classify 98% of subjects correctly in terms of lateralization. The IAP procedure in this study is different from the IAP memory measure in this study, but nonetheless, the high classification accuracy rate further validates this IAP memory procedure.

Specific aim six involved using difference scores to predict post-surgical outcome. Hypothesis 9 involved examining the use of Discrepancy Scores between left and right IAP injection performances on Forms I and II to see if there was a relationship between higher discrepancies (in the expected direction) and subjects who reported complete freedom from disabling seizures (Class I) versus those who continued to report symptoms (Classes II-IV). The results indicated a larger percentage of seizure free subjects (81%) with DS greater than 20% than those non-seizure free subjects with DS greater than 20% (57%) In addition, seizure free subjects had significantly larger mean DS. Consequently, there was a relationship between DS and outcome.

When looking at Form I, there was a significant difference in performance for subjects who were seizure free versus those who were not. Subjects who were not seizure free (Class II-IV) performed significantly worse (mean score = 8.3) on

Form I compared to subjects who were seizure free (mean score = 11.9). This finding may relate to Chelune's (1995) functional reserve model, which asserts that the contralateral temporal lobe to resection is more predictive of patients at risk for developing post surgical memory deficits. Since Form I assessed memory functioning in the contralateral hemisphere to surgical resection, it appears that scores on Form I may be useful in predicting outcome in TLE patients.

Conclusions

There is significant variability in the IAP across epilepsy centers. The IAP memory measures across centers utilize different stimuli and formats to predict post-surgical memory deficits. In addition to variability in procedures, there have been limited studies pertaining to the reliability and validity of IAP memory measures being used, limiting comparisons across centers as to which procedure is the most empirically sound. Consequently, there is a need for individual centers to examine the psychometric properties of their IAP memory measure as well as their success rates for predicting post-surgical memory changes (Simkins-Bullock, 2000).

In this study, the IAP memory measure was examined in terms of its overall psychometric properties. The IAP memory measure demonstrated good reliability across forms. The majority of items on Forms I and II had acceptable item difficulty values, item discrimination values, and item-total correlations.

There were a few items on Forms I and II that may be candidates for revision, but most items were only slightly below predicted ranges. In terms of construct validity, correlations were lower than expected, but provided some evidence of convergent validity for the IAP memory measure. Assessment of convergent validity was likely complicated by the interaction of amobarbital on side of injection, differences in seizure lateralization among subjects, and aphasia during left hemisphere assessment. In terms of divergent validity, both Forms I and II had low correlations with the executive functioning measures, providing some preliminary support for the construct validity of the measure.

When utilizing a discrepancy score to predict lateralization in subjects with TLE, asymmetry scores from the IAP memory measure were able to classify 92% of subjects with either left or right TLE after applying a correction factor for left injection scores. There was limited data regarding post-operative seizure outcome. However, a higher percentage of seizure-free subjects had DS greater than 20% as compared to those subjects who were not seizure-free. In addition, there was a significant difference on Form I between subjects who reported they were seizure-free following temporal resection as compared to subjects who continued to report symptoms.

This study may represent the most detailed examination of the psychometric properties of an IAP procedure to date. Consequently, it is difficult to compare this procedure to others currently used. This IAP measure has several unique characteristics likely to contribute to the reliability of the measure. First, the number of items and foils were based on a binomial probability formula used to calculate the optimal number of foils to minimize the probability of obtaining a "passing score" by random guessing. Second, the IAP memory measure has 18 items (more than most centers) and reliability increases as the number of items on a test increase. Third, the majority of items appear to be testing a similar construct, and this homogeneity of items is directly related to the overall reliability of the measure. In terms of validity, the IAP measure demonstrated convergent and divergent validity with the neuropsychological tests although correlations were lower than expected. The IAP measure also demonstrated clinical utility in that it was able to predict seizure lateralization in 92% of subjects when applying a correction factor. Although there are items that may benefit from revision and improvement, the majority of hypotheses were supported and this IAP memory measure appears to have good overall psychometric properties.

Limitations

Psychometric validation of a test involves several aspects that were not included in this study. Calculation of Cronbach's alpha was the only reliability estimate included in this study. Although there is no reason to suspect differences in scoring the IAP memory measure, interrater reliability was not assessed in this study. In addition, there was no measure of test-retest reliability, as it was impossible and unethical to systematically readminister amobarbital at a later time. Due to the fact that subjects were injected on different sides while administered different forms, there was no measure of alternate forms reliability and no means of direct comparison of the three forms. Another limitation of the study is the small sample. Even though there were 90 patients, when looking at the group of patients by side of injection, the group sample sizes were much smaller. The limited number of subjects in certain analyses necessitates cautious interpretation of some findings. In addition, there were some subjects in which lateralization data, outcome data, and neuropsychological assessment was missing. This limited the comparisons between groups as well as the results regarding outcome.

When Milner and colleagues (1962) extended the use of the IAP to assess memory, the procedure became routinely used to predict future memory decrements. However, it was outside the scope of this study to evaluate how well this measure predicts post-surgical deficits. This is an important limitation of this study and further discussed in terms of future directions.

Directions for Future Research

Replication of the current study using a larger sample size may provide further validation of this measure. A larger sample size would provide an opportunity to conduct factor analysis, which is commonly employed in the validation of psychological tests. A sample size with either 100 subjects or 10 times the number of variables is suggested for an effective factor analysis (Crocker & Algina, 1986). Further validation of this memory measure through correlations with neuroimaging techniques should be examined. The IAP procedure utilized at the Medical College of Georgia had been validated through studies demonstrating the relationship between IAP scores and hippocampal volume asymmetries in subjects with complex partial seizures, and the protocol has been shown to be related to hippocampal pathology and post-operative material-specific memory decline (Loring et al., 1997; Loring et al., 1990; Loring, Murro, Meador, Lee, & et al., 1993; Loring et al., 1994). However, the IAP memory procedure has not been examined in terms of reliability, item-analysis, or convergent validity. It is recommended that subjects undergoing ATL routinely be referred for neuropsychological assessment post-surgery. Predictive validity should then be examined to see if performance on the IAP memory measure is related to material specific deficits assessed by neuropsychological memory measures. Investigation of the relationship between pre and post neuropsychological memory performance and IAP memory testing would help validate this measure with respect to post-surgical memory deficits.

Table 1

Demographic variables and duration of epilepsy

	N	Mean	SD	Range
Age (years)	90	35.6	10.8	18-69
Education (years)	76	12.5	3.2	5-20
FSIQ	57	86.7	14.0	67-127
Duration of epilepsy (years)	73	18.2	12.7	1-50

Note. FSIQ is Estimated Full Scale Intelligence Quotient from the Wechsler Adult Intelligence Scale-Third Edition.

Table 2

Demographic variables in patients with LTLE and RTLE.

Demographic variables/seizure focus				Total
Sex	Males	Females		
RTLE	9	17		26
LTLE	13	25		38
Ethnicity	Caucasian	Other*		
RTLE	14	8		22
LTLE	22	20		42
Handedness	Left	Right	Bilateral	
RTLE	4	18	0	22
LTLE	7	34	1	42
Language Dominance	Left	Right	Mixed	
RTLE	22	0	0	22
LTLE	36	2	4	42

*RTLE = 2 African American, 6 Hispanic. LTLE 7 African American, 10 Hispanic, 1 Asian, 1 Middle Eastern. Chi-Square analyzes indicated no significant differences between sex and ethnicity in RTLE versus LTLE groups. Fisher's Exact Test indicated no significant difference in handedness or language dominance among subjects with RTLE and LTLE.

Table 3

Demographic variables in patients with LTLE and RTLE.

	RTLE	N	Mean	SD	LTLE	N	Mean	SD
Age		22	37.41	12.84		42	36.07	9.74
Education		21	11.81	3.08		33	13.09	3.06
FSIQ		14	88.21	12.42		26	86.65	16.4
Seizure		19	21.71	15.73		34	16.41	11.41

Note. FSIQ Full Scale IQ from Wechsler Adult Intelligence Scale-III. Independent-samples t test indicated no significant differences between RTLE and LTLE groups.

Table 4

Diagnostic Categories and diagnosis by seizure lateralization.

Diagnosis	N	Percent
MTLE	28	31.1
TLE	56	62.2
Nondiagnostic Seizure Evaluation	6	6.7
Total	90	100

Diagnosis Bilateral	N	RTLE	LTLE	
Undetermined seizure focus	21			
MTLE	28	9	17	2
TLE	41	13	25	3
Total	90	22	42	5

MTLE = Mesial Temporal Lobe Epilepsy, TLE = Temporal Lobe Epilepsy.

Table 5

Means and standard deviations for the raw scores on Forms I and II across demographic characteristics.

Form/Variable	N	Mean	SD	N	Mean	SD
Form I	90	10.14	4.47			
Form II	90	7.38	4.47			
Ethnicity	Caucasian			Other		
Form I	52	11.00*	4.36	36	9.08	4.34
Form II	52	8.06	4.57	36	6.47	4.21
Handedness	Right			Left		
Form I	73	9.64	4.63	16	12.25*	3.00
Form II	73	7.29	4.72	16	7.56	3.24
Language	Right			Left		
Form I	4	15.25*	0.96	79	10.05	4.42
Form II	4	6.25	2.75	79	7.70	4.56
Lateralization	Right			Left		
Form I	22	12.50*	3.54	42	8.93	4.27
Form II	22	4.32	2.66	42	8.95*	4.45

*Indicates significantly better performance on that Form. $p < .05$. Independent-samples t tests were used to compare mean performance for variables on each form. Ethnicity refers to Caucasians versus all other ethnicities. Language refers to side of language dominance. Lateralization refers to side of seizure focus.

Table 6

Calculation of Cronbach's Alpha for IAP memory measure if item is deleted across Form I, II, and III

Form I	r	Form II	r	Form III	r
1	.83	1	.82	1	.70
2	.83	2	.84	2	.70
3	.84	3	.83	3	.69
4	.83	4	.83	4	.66
5	.85	5	.83	5	.69
6	.84	6	.82	6	.66
7	.83	7	.83	7*	
8	.84	8	.83	8	.66
9	.84	9	.82	9	.69
10	.84	10	.82	10	.65
11	.84	11	.83	11	.68
12	.84	12	.82	12	.65

Table 6 Continued

13	.84	13	.82	13	.69
14	.85	14	.82	14	.67
15	.84	15	.83	15	.67
16	.84	16	.82	16	.67
17	.84	17	.82	17	.70
18	.83	18	.83	18	.68

*Item 7 had zero variance and was removed from analysis. Note. This table examines whether each the reliability for each Form increases or decrease if individual item is deleted.

Table 7

Mean item difficulty values by side of injection

Form	Side of injection	Mean	N
Form I	Right	.68	33
	Left	.50	57
Form II	Right	.47	57
	Left	.31	33

Table 8

Item Difficulty for all subjects and across second injection side for Form II

	For all subjects	Right 2nd injection (N=57)	Left 2nd injection (N=33)
1.	.73	.91	.63
2.	.38	.58	.26
3.	.60	.64	.58
4.	.51	.73	.39
5.	.44	.39	.47
6.	.49	.70	.37
7.	.69	.73	.67
8.	.37	.52	.28
9.	.71	.70	.72
10.	.46	.61	.37
11.	.24	.30	.21
12.	.43	.64	.32
13.	.84	.85	.84
14.	.41	.55	.33

Table 8 Continued

15.	.81	.91	.75
16.	.61	.82	.49
17.	.82	.91	.77
18.	.64	.85	.53

Note. Picture items are odd numbered items and word items are even numbered items.

Table 9

Item Difficulty for all subjects and across second injection side for Form II

	For all subjects	Right 2nd injection (N=57)	Left 2nd injection (N=33)
1.	.41	.51	.24
2.	.52	.53	.52
3.	.33	.32	.36
4.	.33	.39	.24
5.	.12	.12	.12
6.	.33	.44	.15
7.	.64	.67	.60
8.	.43	.51	.30
9.	.49	.54	.39
10.	.28	.37	.12
11.	.44	.49	.36
12.	.32	.44	.12
13.	.46	.47	.42
14.	.36	.47	.15

Table 9 Continued

15.	.33	.23	.52
16.	.42	.56	.18
17.	.60	.70	.42
18.	.60	.70	.42

Note. Picture items are odd numbered items and word items are even numbered items.

Table 10

Item Difficulty for all subjects on Form III

1.	.99
2.	.86
3.	.98
4.	.87
5.	.92
6.	.49
7.	1.00
8.	.79
9.	.98
10.	.63
11.	.88
12.	.83
13.	.94
14.	.74
15.	.88
16.	.62
17.	.96
18.	.97

Note. Picture items are odd numbered items and word items are even numbered items.

Table 11

Item discrimination for Forms I-III

Item Number	Form I	Form II	Form III
1.	.60	.82	.02
2.	.67	.43	.14
3.	.53	.42	.05
4.	.74	.32	.30
5.	.35	.32	.13
6.	.59	.61	.49
7.	.69	.57	0
8.	.50	.37	.33
9.	.58	.71	0
10.	.46	.56	.57
11.	.28	.44	.24
12.	.62	.59	.30
13.	.28	.62	.08
14.	.22	.54	.34

Table 11 Continued

15.	.49	.33	.24
16.	.53	.69	.56
17.	.44	.83	.05
18.	.63	.41	.07

Table 12

Item-Total Correlations for Forms I, II, and III

Item Number	Form I	Form II	Form III
1.	.55	.54	-.03
2.	.65	.17	.06
3.	.44	.38	.15
4.	.53	.34	.40
5.	.23	.23	.18
6.	.46	.54	.40
7.	.58	.41	0
8.	.44	.29	.45
9.	.45	.56	.12
10.	.42	.56	.48
11.	.34	.29	.26
12.	.46	.58	.48
13.	.33	.48	.10
14.	.22	.48	.37
15.	.46	.35	.37
16.	.45	.55	.37
17.	.49	.58	-.01
18.	.50	.33	.36

Table 13

Correlations between Forms I and II and neuropsychological measures for all subjects.

Memory Test	N	Form I	Form II
		Total Score	Total Score
LM II	75	.30**	.21
CVLT	70	.39**	.13
VR II	72	.23	.31**
ROCF	76	.12	.36**
		Word Score	Word Score
LM II	76	.28*	.16
CVLT	71	.39*	.06
VR II	73	.13	.21
ROCF	76	-.01	.31*
		Picture score	Picture score
LM II	76	.22	.25*
CVLT	71	.26*	.21
VR II	73	.28*	.36**
ROCF	76	.22	.34*

LM II = Logical Memory II, VR II =Visual Reproduction II, CVLT = California Verbal Learning Test. CVLT score is the total score for the long delay free recall, Rey-Osterrieth Complex Figure (ROCF) score is the total score for delayed recall. $^{**}P < 0.01$. $^{*}P < 0.05$.

Table 14

Performance on IAP Forms and Neuropsychological testing for all subjects.

Memory Test Scores	N	Mean	SD
Form I total score	90	10.14	4.47
Form I word total	90	4.30	2.65
Form I picture total	90	5.90	2.28
Form II total score	90	7.38	4.47
Form II word total	90	3.60	2.48
Form II picture total	90	3.83	2.48
Form III total score	90	15.16	2.80
LM II Raw score	76	18.88	10.17
VR II Raw score	73	41.68	23.02
CVLT LD Free recall	71	8.61	3.62
Rey-O Delay Delayed recall	76	15.25	6.39
TMT A (Seconds)	76	34.88	15.30
TMT B (Seconds)	76	97.57	46.23
WCST #PSV	69	24.22	19.21

LM II = Logical Memory II, VR II = Visual Reproduction II, CVLT = California Verbal Learning Test. CVLT score is the total score for the long delay free recall. TMT A = Trail Making Test Part A, TMT B = Trail Making Test Part B, WCST #PSV = Wisconsin Card Sorting Test number of perseverations.

Table 15

Performance on IAP Forms and neuropsychological measures for TLE subjects injected on the right side first.

Memory Test Scores	N	Mean	SD
Form I total score	31	12.33	3.93
Form I word total	31	5.97	2.27
Form I picture total	31	6.33	1.96
Form II total score	31	5.45	3.49
Form II word total	31	2.21	1.58
Form II picture total	31	3.45	2.65
Form III total score	31	14.58	3.58
LM II Raw score	29	19.94	10.78
VR II Raw score	29	41.58	20.57
CVLT LD Free recall	29	9.23	3.22
Rey-O Delayed recall	29	15.03	5.93
TMT A (Seconds)	29	32.26	13.31
TMT B (Seconds)	29	101.26	55.45
WCST # PSV	26	23.15	18.75

LM II = Logical Memory II, VR II = Visual Reproduction II CVLT = California Verbal Learning Test. CVLT score is the total score for the long delay free recall. TMT A = Trail Making Test Part A, TMT B = Trail Making Test Part B, WCST #PSV = Wisconsin Card Sorting Test number of perseverations.

Table 16

Correlations between cognitive measures and IAP scores for subjects with TLE injected on the right side first and left side second.

Memory Test	N	Form I (Left Memory)	Form II (Right Memory)
LM II	29	.57**	.09
CVLT	29	.49**	.30
VR II	29	.29	.42*
ROCF	29	.21	.29
		Word Score	Word Score
LM II	29	.53**	.11
CVLT	29	.56**	.12
VR II	29	.21	.15
ROCF	28	.07	.04
		Picture score	Picture score
LM II	29	.44*	.12
CVLT	29	.25	.36
VR II	29	.29	.49**
ROCF	29	.22	.33

LM II =Logical Memory II, VR II =Visual Reproduction II, CVLT =California Verbal Learning Test. CVLT score is the total score for the long delay free recall, the Rey-Osterrieth Complex Figure (ROCF) score is the total score for delayed recall. **$P < 0.01$. *$P < 0.05$.

Table 17

Performance on IAP Forms and neuropsychological measures for subjects with TLE injected on the left side first.

Memory Test Scores	N	Mean	SD
Form I total score	53	8.88	4.30
Form I word total	53	3.33	2.37
Form I picture total	53	5.65	2.42
Form II score	53	8.49	4.61
Form II word total	53	4.40	2.56
Form II picture total	53	4.05	2.37
Form III total score	53	15.49	2.20
LM II Raw score	42	18.16	9.79
VR II Raw score	42	41.76	24.92
CVLT LD Free recall	42	8.13	3.86
Rey-O Delayed recall	42	15.39	6.74
TMT A (Seconds)	42	36.69	16.44
TMT B (Seconds)	42	95.02	39.12
WCST #PSV	42	24.86	19.67

LM II = Logical Memory II, VR II = Visual Reproduction II, CVLT = California Verbal Learning Test. CVLT score is the total score for the long delay free recall. TMT A = Trail Making Test Part A, TMT B = Trail Making Test Part B, WCST #PSV = Wisconsin Card Sorting Test number of perseverations.

Table 18

Correlations between cognitive measures and IAP scores for TLE subjects injected on the left side first.

Memory Test	N	Form I (Right Memory)	Form II (Left Memory)
LM II	42	.08	.23
CVLT	42	.28	.26
VR II	42	.19	.23
ROCF	42	.10	.34*
		Word Score	Word Score
LM II	42	.08	.22
CVLT	42	.25	.30
VR II	42	.07	.20
ROCF	42	-.04	.37*
		Picture score	Picture score
LM II	42	.04	.23
CVLT	42	.22	.20
VR II	42	.24	.24
ROCF	42	.24	.27

LM II =Logical Memory II, VR II =Visual Reproduction II, CVLT =California Verbal Learning Test. CVLT score is the total score for the long delay free recall, the Rey-Osterrieth Complex Figure (ROCF) score is the total score for delayed recall. $^{**}P < 0.01$. $^{*}P < 0.05$.

Table 19

Performance on IAP Forms and neuropsychological measures for subjects with RTLE.

Memory Test Scores	N	Mean	SD
Form I total score	22	12.5	3.54
Form I word total	22	6.05	2.01
Form I picture total	22	6.45	1.87
Form II total score	22	4.32	2.66
Form II word total	22	1.73	1.12
Form II picture total	22	2.68	2.48
Form III total score	22	15.00	2.78
LM II Raw score	21	21.14	9.70
VR II Raw score	21	40.86	21.12
CVLT LD Free recall	21	9.43	3.44
Rey-O Delayed recall	21	15.19	6.55
TMT A (Seconds)	21	33.71	14.13
TMT B (Seconds)	21	104.19	62.88
WCST #PSV	18	23.61	17.83

LM II = Logical Memory II, VR II = Visual Reproduction II, CVLT = California Verbal Learning Test. CVLT score is the total score for the long delay free recall. TMT A = Trail Making Test Part A, TMT B = Trail Making Test Part B, WCST #PSV = Wisconsin Card Sorting Test number of perseverations.

Table 20

Correlations between cognitive measures and IAP scores for subjects with RTLE.

Memory Test	N	Form I (Left Memory)	Form II (Right Memory)
LM II	21	.55**	.26
CVLT	21	.45*	.48*
VR II	21	.35	.44*
ROCF	21	.28	.37
		Word Score	Word Score
LM II	21	.52*	.22
CVLT	21	.53*	.23
VR II	21	.30	.11
ROCF	21	.14	.10
		Picture score	Picture score
LM II	21	.36	.26
CVLT	21	.17	.49*
VR II	21	.28	.50*
ROCF	21	.24	.37

LM II =Logical Memory II, VR II =Visual Reproduction II, CVLT =California Verbal Learning Test. CVLT score is the total score for the long delay free recall, the Rey-Osterrieth Complex Figure (ROCF) score is the total score for delayed recall. $^{**}P < 0.01$. $^{*}P < 0.05$.

Table 21

Performance on IAP Forms and neuropsychological measures for subjects with LTLE.

Memory Test Scores	N	Mean	SD
Form I total score	42	8.8	4.22
Form I word total	42	3.20	2.28
Form I picture total	42	5.70	2.52
Form II total score	42	9.05	4.52
Form II word total score	42	4.65	2.62
Form II picture total score	42	4.38	2.26
Form III total score	42	15.45	2.18
LM II Raw score	33	18.24	10.02
VR II Raw score	31	44.16	22.98
CVLT LD Free recall	30	8.37	4.15
Rey-O Delayed recall	33	15.70	7.35
TMT A (Seconds)	32	35.22	14.10
TMT B (Seconds)	32	92.75	42.27
WCST #PSV	31	24.26	20.98

LM II = Logical Memory II, VR II = Visual Reproduction II, CVLT = California Verbal Learning Test. CVLT score is the total score for the long delay free recall. TMT A = Trail Making Test Part A, TMT B = Trail Making Test Part B, WCST #PSV = Wisconsin Card Sorting Test number of perseverations.

Table 22

Correlations between memory measures and total score for subjects with LTLE.

Memory Test	N	Form I (Right Memory)	Form II (Left Memory)
LM II	33	.09	.37*
CVLT	30	.28	.30
VR II	31	.14	.26
ROCF	33	.13	.35*
		Word Score	Word Score
LM II	33	.06	.39*
CVLT	30	.25	.36*
VR II	31	.02	.23
ROCF	33	-.03	.39*
		Picture score	Picture score
LM II	33	.09	.30
CVLT	30	.23	.19
VR II	31	.20	.26
ROCF	33	.27	.25

LM II =Logical Memory II, VR II =Visual Reproduction II, CVLT =California Verbal Learning Test. CVLT score is the total score for the long delay free recall, the Rey-Osterrieth Complex Figure (ROCF) score is the total score for delayed recall. $**P < 0.01$. $*P < 0.05$.

Table 23

Discrepancy Scores for Patients with TLE for use in logistic regression.

Diagnosis	N	Min	Max	Mean	SD
RTLE	22	22.22	66.67	45.45	15.30
Corrected RTLE	22	32.22	76.67	55.45	15.30
LTLE	42	-61.11	66.67	1.11	29.75
Corrected LTLE	42	-71.11	56.67	-8.61	29.74

Note. N is the total number of subjects with left or right TLE. DS Discrepancy scores $\geq$ 20% suggested significant asymmetry with negative scores indicating better performance in the right hemisphere and positive scores indicating better performance in the left hemisphere. DS less than 20% were considered nonlateralized or indeterminate. Correction was 10 percentage points added to DS. Corrected RTLE = Corrected score for left injection in subjects with RTLE. Corrected LTLE = Corrected score for left injection in subjects with LTLE.

Appendix A

The California Verbal Learning Test (CVLT):

The CVLT (Delis et al., 1987) is an individually-administered measure of verbal learning and memory. Sixteen words (Monday list) are presented orally, and subjects are asked to repeat as many items as they can, in any order, for each of five (learning) trials. The words are from four categories (fruits, tools, clothing, spices and herbs), although they are not organized as such in their presentation to the subject. After the fifth trial, subjects are asked to repeat words from a new list of sixteen items (Tuesday list). Immediately after the Tuesday list, subjects are asked to recall as many items as they can from the Monday list (short delay free recall), and then are cued to recall the items from each of the four categories (short delay cued recall). After approximately 20 minutes, the recall procedure for the Monday list is repeated for both the free recall (long delay free recall), and cued recall (long delay cued recall) conditions. Next, a longer list of words comprised of the entire Monday list, parts of the Tuesday list and several novel words is presented, and subjects are asked to identify which words were part of the Monday list (recognition). Computerized scoring yields standard scores for total learning, all four recall conditions, and recognition. Also included are learning characteristics, such as whether the subjects encoded the list in order of the words presented (serial clustering), or by grouping words into four categories (semantic clustering).

Internal consistency reliability coefficients for the CVLT range from r = .69 to r =.92 (Delis et al., 1987), and factor analyses yielded a six factor solution, including general verbal learning, response discrimination, learning strategy, proactive effect, serial position effect, and acquisition rate (Delis, Freeland, Kramer & Kaplan, 1988).

Rey-Osterrieth Complex Figure Test (Rey-O)

The Rey-Osterrieth Complex Figure Test (Loring, Martin, Meador, & Lee, 1990; Osterrieth, 1944) is a measure of visuospatial constructional ability and visual memory. There are four components in the version of the Rey-O used in this study: Copy, Immediate Recall, Delayed Recall, and Recognition. The subject is given a blank sheet of paper (8 ½'' by 11") and asked to reproduce the figure as it is drawn on the stimulus sheet, using three colored pencils given to the subject by the examiner at roughly equal points of the subject's reproduction. Afterwards, the stimulus and drawing are removed and the subject is asked to reproduce the figure from memory, using the same three-pencil method. Following a 15 minute delay, the subject is again asked to reproduce the figure from memory. Finally, the subject is asked to choose the original figure from eight similar-looking drawings during the recognition portion of the test. Standardized scoring of the protocols is used (Loring et al., 1990). Inter-rater reliability coefficients are high for all three trials, r = .80 for the copy, r =.93 for immediate recall, and r =.96 for delayed recall (Loring, Martin, Meador, & Lee, 1990). One year test-retest reliability coefficients range from r =.47 to r =.59 for immediate and delayed recall trials (Berry, Allen, & Schmitt, 1991).

Wechsler Memory Scale-III (WMS-III)

Two subtests of the WMS-III (Wechsler, 1997), Logical Memory and Visual Reproduction, were examined in this study. Logical Memory assesses immediate and delayed recall of structured verbal information in the form of two short stories. The subject is read the stories and then asked to recall as much of it he or she can remember immediately after the stories are read and again after a 30-minute delay. The second story is read twice to the subject. Visual Reproduction assesses immediate and delayed recall of nonverbal information in the form of five cards with seven simple geometric figures. The subject is presented each stimulus, one at a time for 10 seconds, and then asked to draw the stimulus immediately from memory and again after a 30 minute delay. Scoring of these measures is rather detailed, and complete rules are outlined in the WMS-III manual. The WMS-III was co-normed with the Weschler Adult Intelligence Scale-Third Edition (WAIS-III) according to the 1995 U.S. census data. Internal consistency assessed using split-half reliability estimates for Logical Memory and Visual Reproduction ranged from r =. 77 to r = .88. Inter-rater reliability coefficients for Logical Memory and Visual Reproduction were greater than r = .90, using the scoring rules defined in the manual (Wechsler, 1997).

Wisconsin Card Sorting Test (WCST)

The WCST (Heaton et al., 1981) is a measure of flexibility of thought and problem solving. Examinees are asked to match response cards each to one of four stimulus cards using the sorting principles of color (red, green, yellow, blue), form (triangles, stars, crosses, circles), or number (one, two, three, four). These principles are not explicitly explained, although correct/incorrect feedback is given for each response. After ten correct matches in a row, the correct sorting principle is covertly changed. The test continues until the patient has made six runs of ten consecutive correct matches or has not completed any categories after 64 card placements, or has used all 128 response cards. The WCST yields scores such as number of perseverative responses, loss of set, and number of categories completed.

The WCST has excellent inter-rater reliability, with coefficients of r =.93 for perseverative responses, r =.92 for perseverative errors and r = .88 for non-perseverative errors (Axelrod, Goldman, & Woodard, 1992). Extensive research on the concurrent and construct validity of the WCST has been conducted, with most results suggesting the WCST is a valid measure of “executive functioning” (See WCST manual).

Trail Making Tests, Parts A (TMT A) & B (TMT B)

The Trail Making Tests (Reitan, 1955) provide a measure of attention, scanning speed, mental sequencing, and flexibility of thought (Spreen et al., 1988). TMT A requires drawing a line to sequentially connect encircled numbers (1 to 25) on a page. During TMTB, subjects alternate between drawing lines from encircled numbers to encircled letters (e.g., 1-A-2-B-3-C etc. to 13). The reader may consult Spreen and Struass (1988) for a complete description of standardized administration procedures. Raw scores for time to complete the two conditions may be converted to age, sex, and education-corrected T scores. The test has been shown to be sensitive to neuropsychological dysfunction in a wide variety of individuals (Lezak, 1995).

Appendix B

Form I Answers

Wada I Recall Score: Rinj Linj Baseline Name:____________ Date:________ Score:________

	1	2	3	4	5	6	7	8
1	horse	polar bear	bull	lion	**cow**	moose	elephant	rabbit
2	he	**mine**	them	us	hers	ourselves	they	those
3	stick tennis	stick baseball	stick rowing	stick volleyball	stick field hockey	stick soccer	**stick cycling**	stick shooting
4	headlight	hillside	hotel	**hammer**	ham	house	history	hatchet
5	MA	HI	MI	AK	**WA**	OK	TX	NY
6	before	include	**except**	after	without	down	exclude	since
7	football	volleyball	**life preserver**	darts	bat and ball	hockey puck and stick	soccer ball	basketball and hoop
8	pyramid	raptor	snake	volcano	funnel	**river**	acorn	roach
9	4 rings	2 rings	6 arrows	**4 curves**	4 arrows	8 arrows	4 fat arrows	4 + arrows
10	low	aside	despite	**though**	under	beneath	into	along
11	Australia	Canada	Greece	Mexico	Egypt	Brazil	India	**China**
12	**lantern**	igloo	cactus	elevator	dominoes	saguaro	scissors	broom
13	basketball player	tennis player	**gymnast**	football player	baseball player	high jumper	hurdler	runner
14	**between**	toward	instead	divide	often	beyond	inspire	neither
15	5 point star flag	diagonal bar flag	**4 point star flag**	5 stripe & triangle flag	standing bear flag	lion flag	sword & laurel flag	abstract chevron flag
16	table	harp	nail	ticket	needle	**pajamas**	tomb	tabloid
17	Sydney opera house	Sphinx	**Taj Mahal**	Cathedral	Branden-berg gate	Great Wall of China	Greek temple	Mount Rushmore
18	Humpty Dumpty	**Jack Sprat**	Mary had a little lamb	Jack and Jill	Jack be nimble	Mary Mary quite	Three blind mice	Hickory dickory dock

Form II Answers

Wada II Recall Score: Rinj Linj Baseline Name:__________ Date:_______ Score:_______

	1	2	3	4	5	6	7	8
1	hammer	safety goggles	handsaw	toolbox	pliers	vise	screws	steel jaw trap
2	her	yours	she	me	him	these	its	we
3	stick boxer	stick fencing	stick weight lifter	stick roller skater	stick ice hockey	stick speed skating	stick ice skating	stick jockey
4	alligator	auto	apple	ascot	arrow	amplifier	angel	anteater
5	VA	FL	LA	CA	AZ	MD	ME	MT
6	need	indeed	much	more	destroy	alongside	really	forgone
7	scale	coffee cup	light bulbs	fan	lock and keys	battery	measure cup	trash can and lid
8	octopus	hanger	bed	tree	pencil	duplex	trophy	freighter
9	block of 4 rectangle	block of 3 triangle	6 stacked rectangle	2 tubes in perspectiv e	pyramid of stacked layers	4 pyramids	4 receding rectangles	2 receding rectangles
10	moreover	form	Den	rest	inside	regard	besides	emerge
11	South America	North America	Southeast Asia	Europe	Africa	Asia	Central America	Middle East
12	canoe	wreath	feather	globe	galaxy	oboe	Jupiter	unicorn
13	$	Pts	NT$	¥	£	Bs	RbI	DM
14	among	therein	just	such	during	mean	below	while
15	13 star rectangle dark field	3 star and 2 stripe	13 star light wide field	13 star dark field odd stripe	St George cross and stripes	Betsy Ross flag	straight snake and stripes	coiled snake no stripes
16	hammock	pelican	sparrow	monkey	mosquito	eagle	lounge	rocker
17	Orange	Daffodil	Artichoke	Cherries	Lilly	Banana	Strawberr y	Tulips
18	Old MacDonal d	Simple Simon met a pieman	Baa, Baa black sheep	Twinkle twinkle little star	This old man, he played one	Little Tommy Tucker	Quoth the raven, nevermore	Rain rain go away

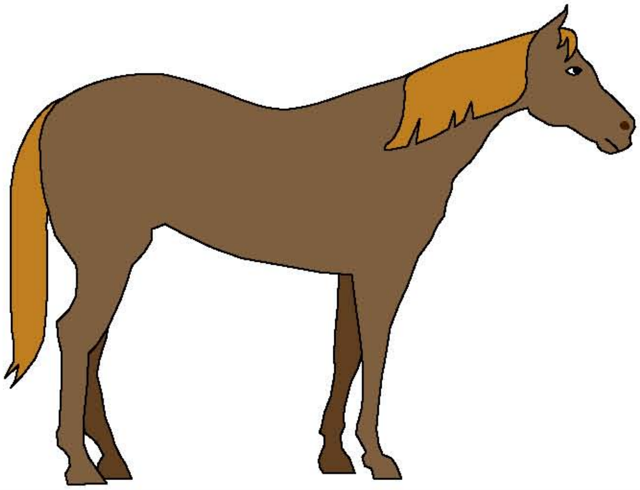
Form III Answers

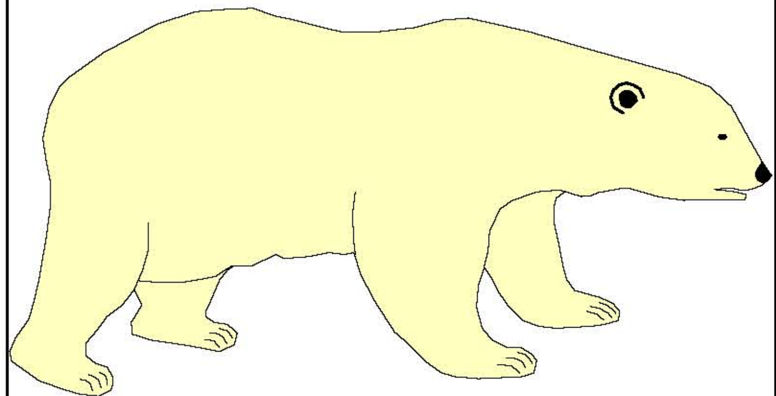
Wada III Recall Score: Rinj Linj Baseline Name:__________ Date:_______ Score:_______

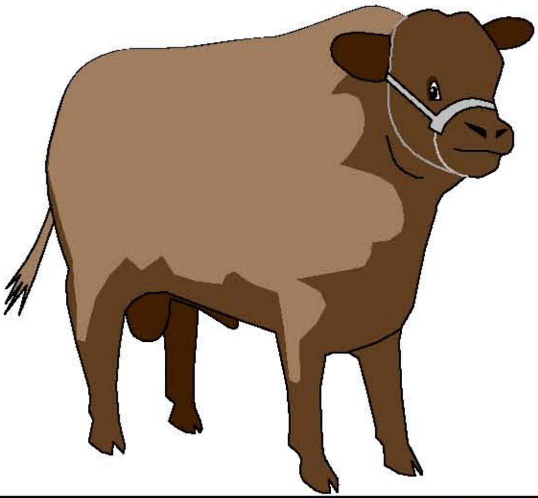
	1	2	3	4	5	6	7	8
1	dish soap dispenser	coffee maker	mailbox	shopping cart	micro-wave oven	clock	lamp	lockbox
2	myself	you	them	his	theirs	it	herself	ours
3	martini	phone sign	police officer sign	help desk sign	sleeping in bed sign	men and women restroom	envelope sign	fork and knife sign
4	cat	chair	car	coin	column	cowboy	camper	cigar
5	AL	CT	NV	IN	IL	ID	MN	UT
6	justify	behind	around	latter	unless	beside	after	over
7	hot-dog	wheat	corn	cheese	apple	grape cluster	carrot	cotton
8	phone	whistle	latch	comb	airplane	compass	rope	daisy
9	4 arrows in circle	6 teardrops in circle	8 curved arrows and circle	thick circular arrow	many circles in circle	3 arrows in circle	cubes and 3D arrows	8 thick arrows in circle
10	usual	gather	about	further	alike	forming	rest	retrieve
11	Denmark	Sweden	Spain	England	Palestine	France	Wales	Italy
12	pepper	fork	frog	piano	toast	fern	potato	photo
13	calendar	hand bones	teeth	cloud	skull	tooth	letter	picket fence
14	hollow	create	restore	wash	point	upon	rule	compress
15	black star rt facing bear	dark star bear right & mission	Texas state flag	7 star and 3 strip flag	eagle and stripes flag	eagle and triangles	flag of president	the people's rights flag
16	journal	meadow	beaver	violin	giraffe	hyena	library	magazine
17	Sun hidden by Cloud	Sun	Lightning	Micropho ne and cloud	Snow and rain	Snow	Sprinkles	Cloud
18	Peter Peter pumpkin eater	Roses are red, violets are blue	One Two buckle my shoe	Ring around the rosy	To be or not to be	Little Miss Muffit	Four and twenty blackbirds	California here I come

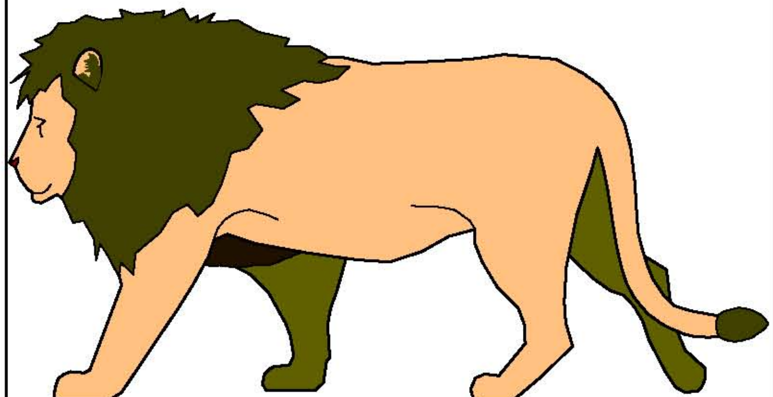
Revised 8-10-99

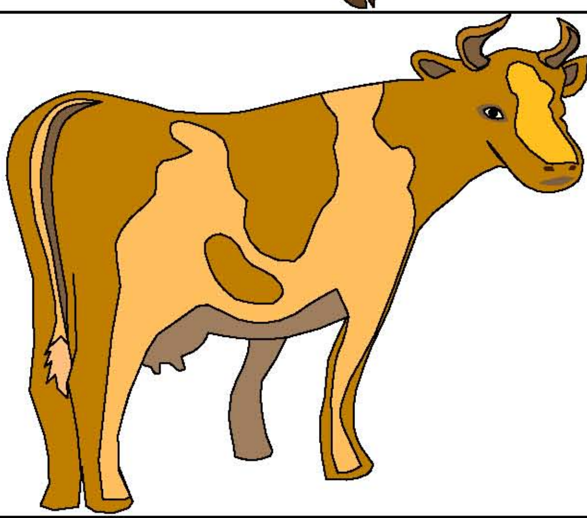
Wada I English Answers

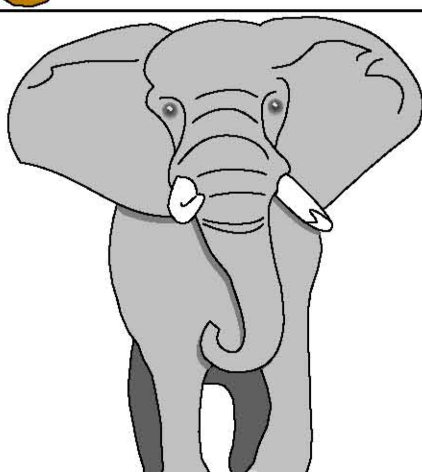


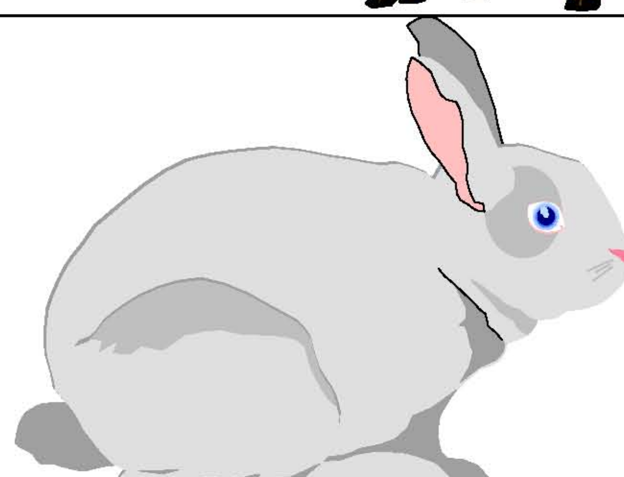




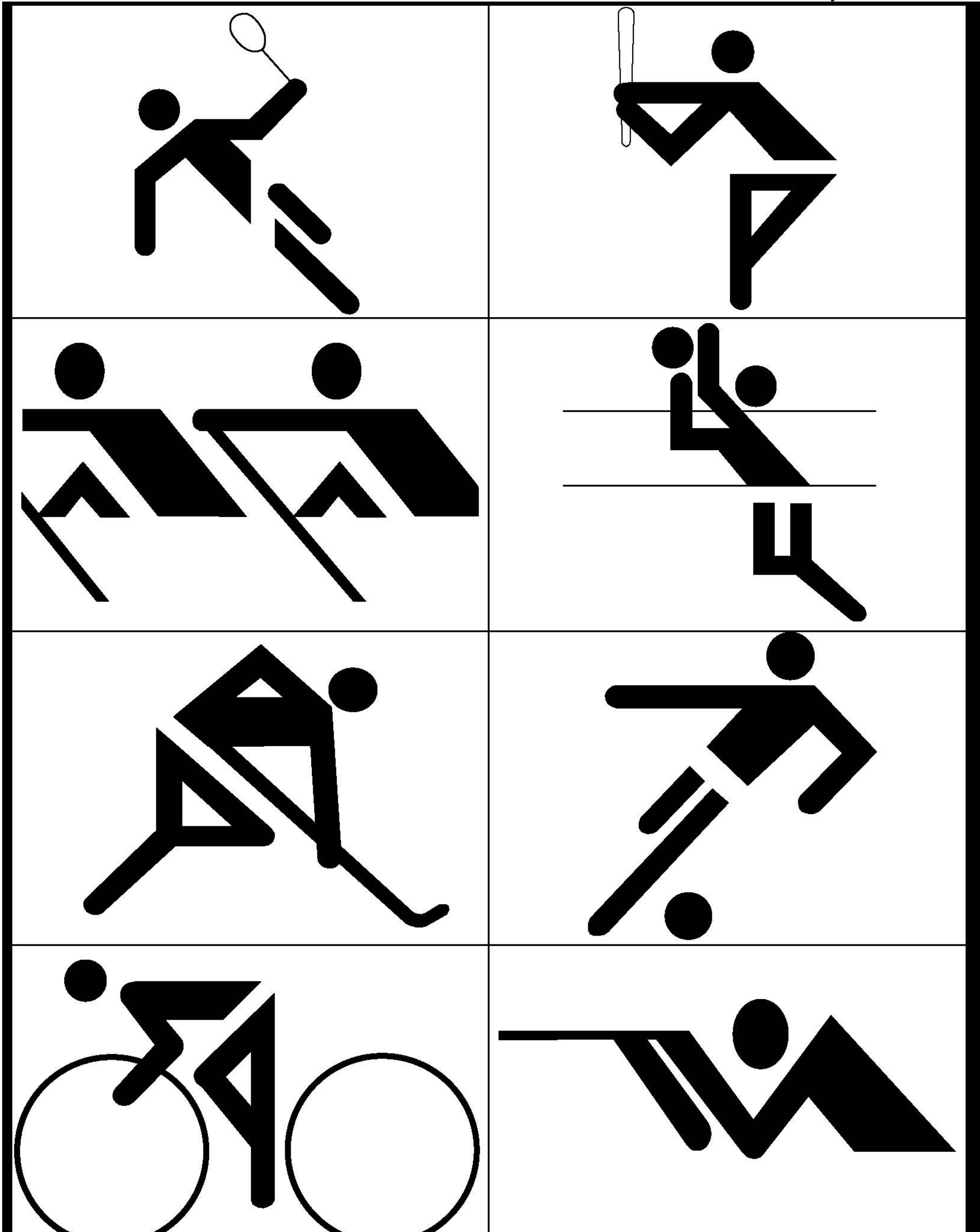




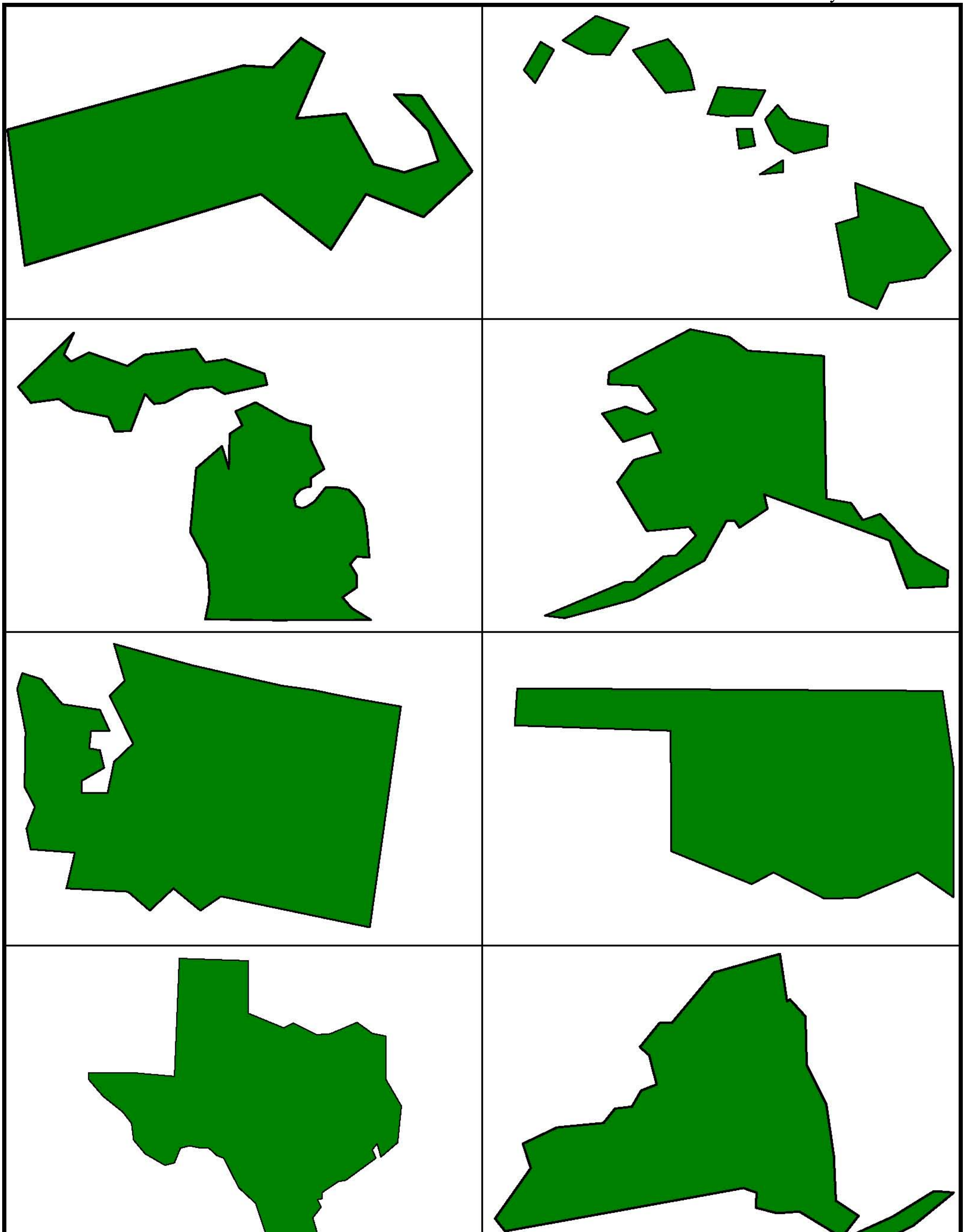




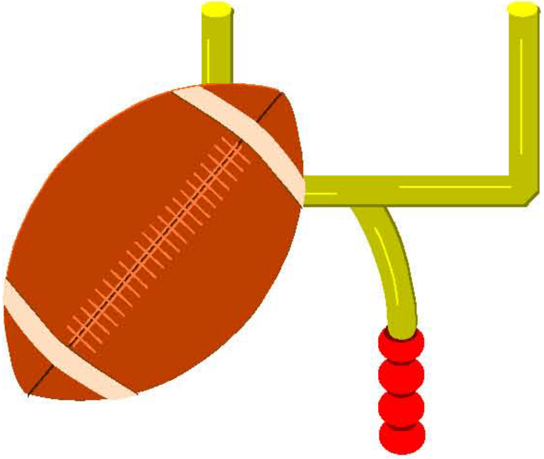
HE	MINE
THEM	US
HERS	OURSELVES
THEY	THOSE

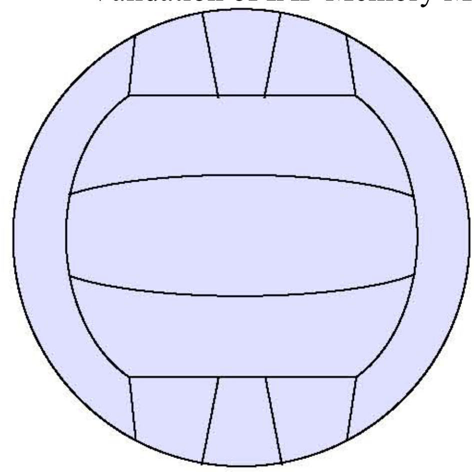


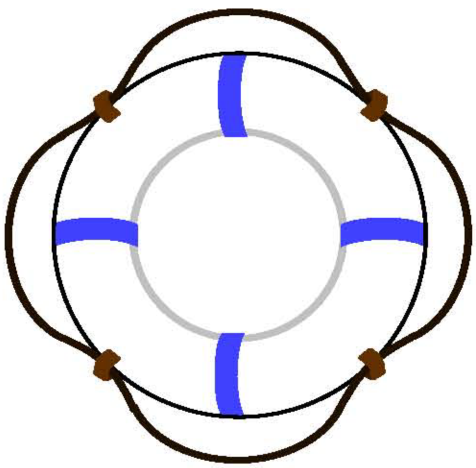
HEADLIGHT	HILLSIDE
HOTEL	HAMMER
HAM	HOUSE
HISTORY	HATCHET

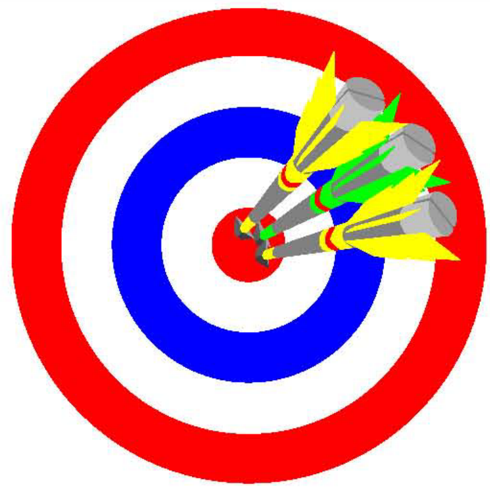


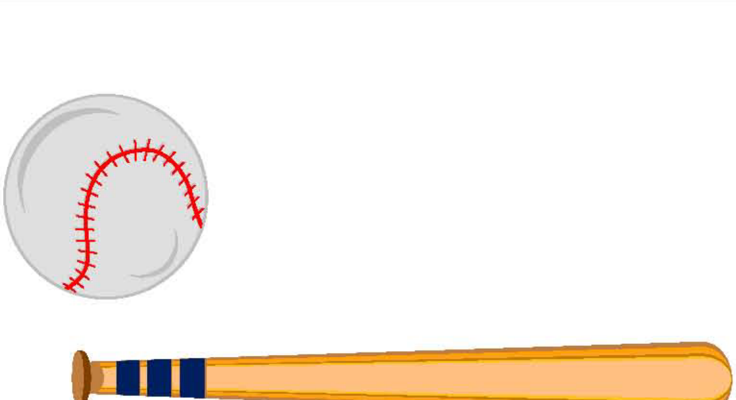
BEFORE	INCLUDE
EXCEPT	AFTER
WITHOUT	DOWN
EXCLUDE	SINCE

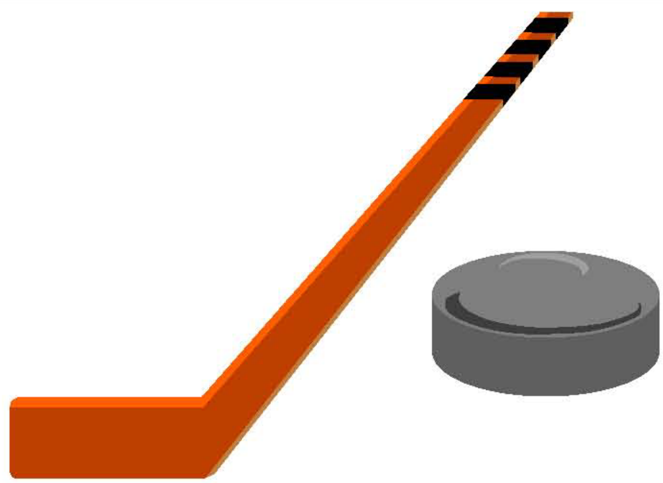


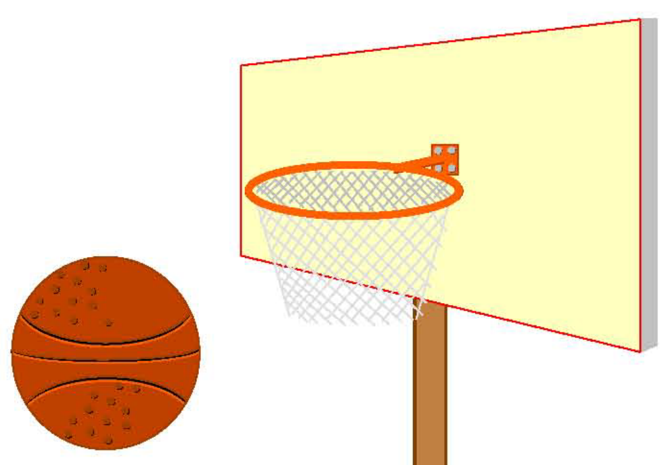




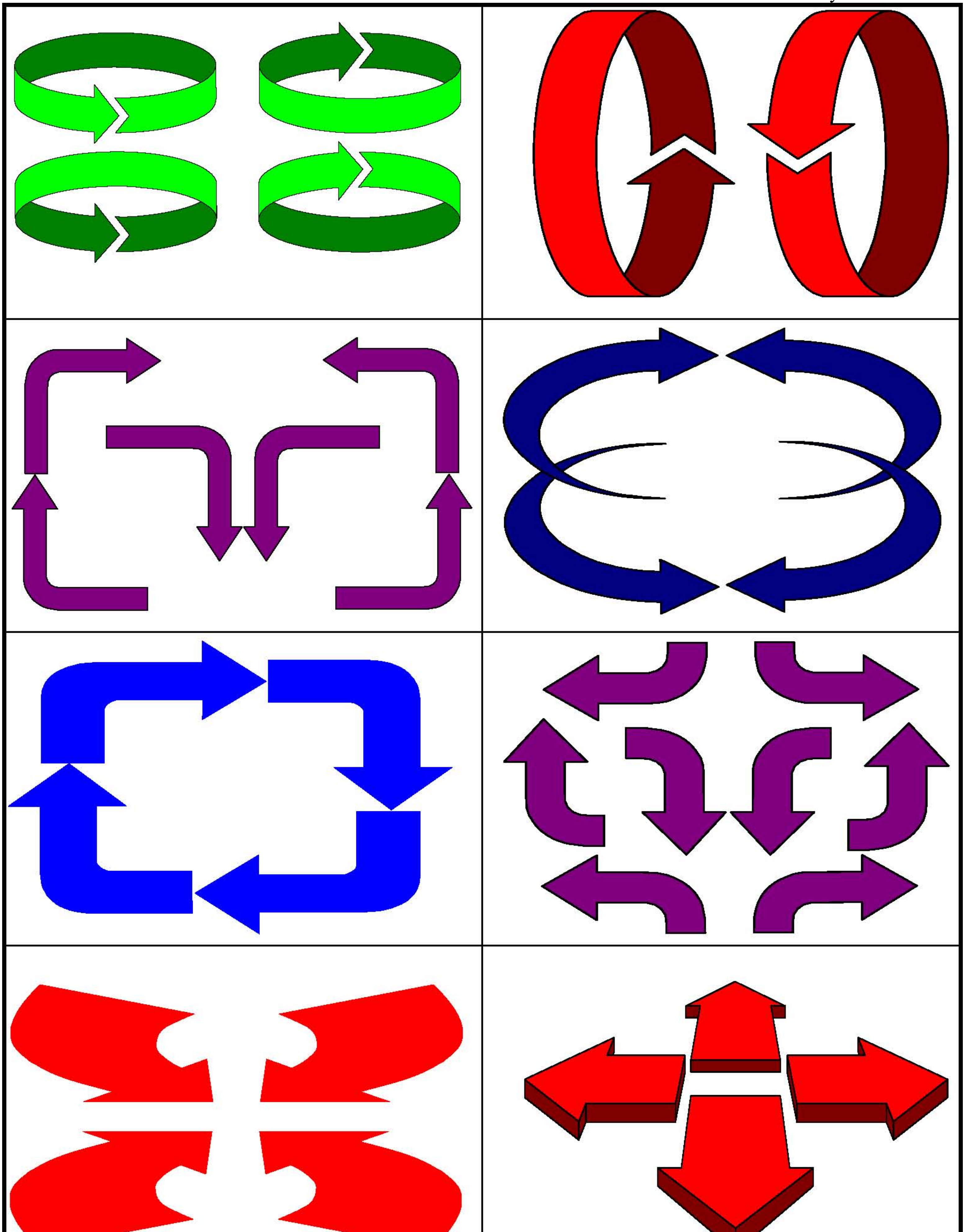






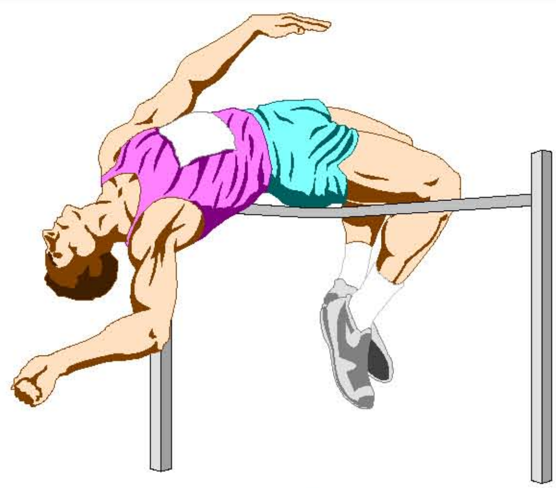


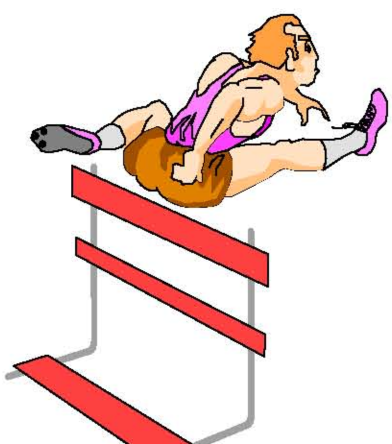
PYRAMID	RAPTOR
SNAKE	VOLCANO
FUNNEL	RIVER
ACORN	ROACH



LOW	ASIDE
DESPITE	THOUGH
UNDER	BENEATH
INTO	ALONG

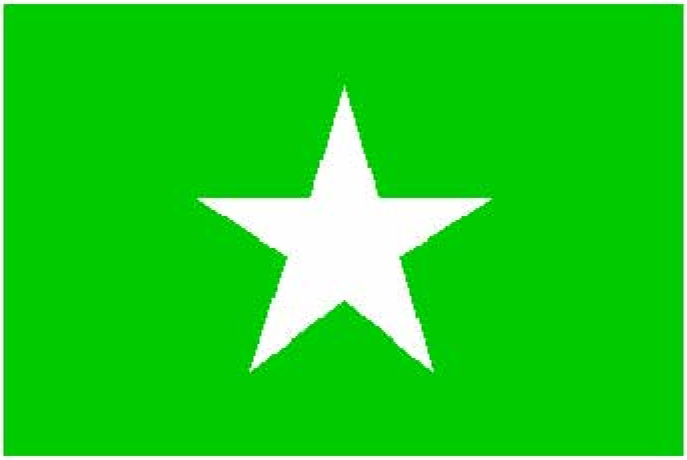
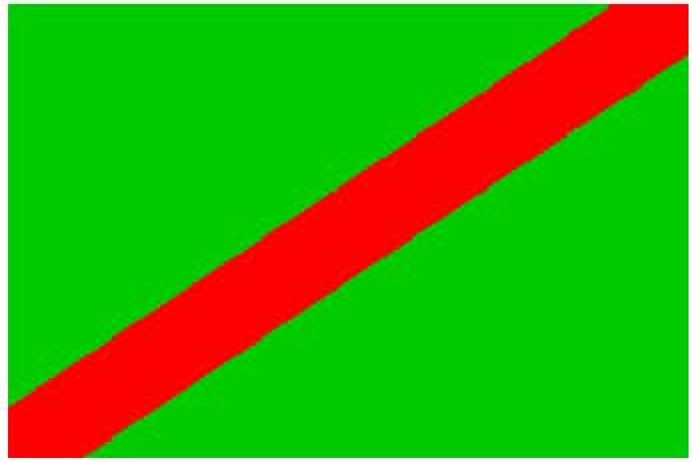
LANTERN	IGLOO
CACTUS	ELEVATOR
DOMINOES	SAGUARO
SCISSORS	BROOM

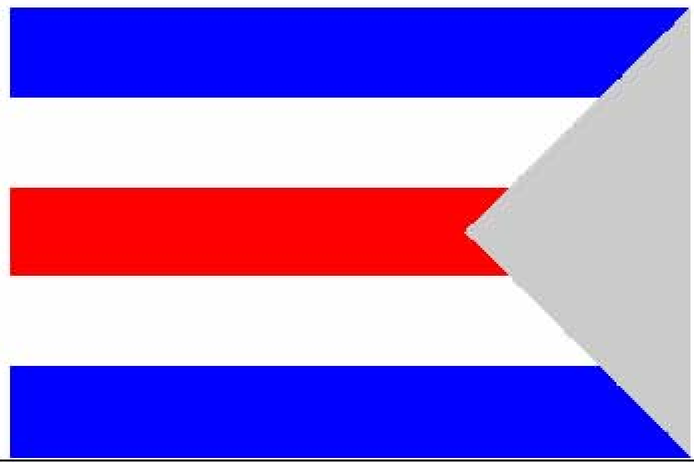





111

BETWEEN	TOWARD
INSTEAD	DIVIDE
OFTEN	BEYOND
INSPIRE	NEITHER





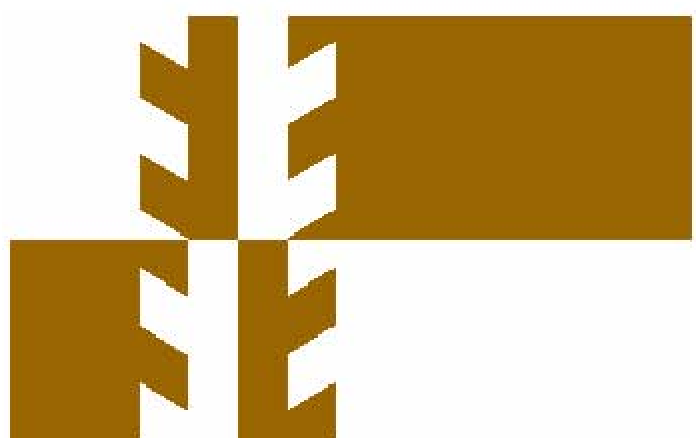
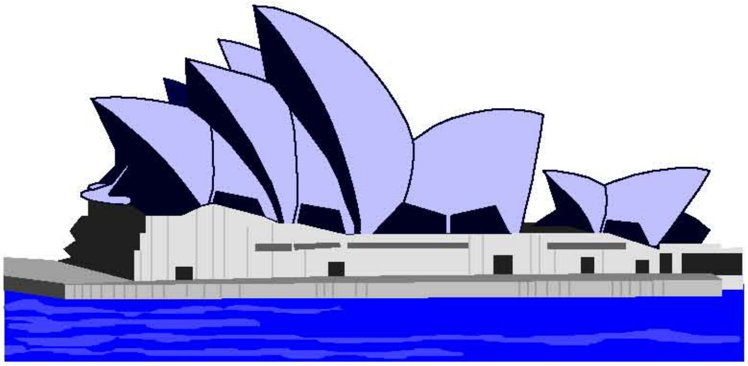
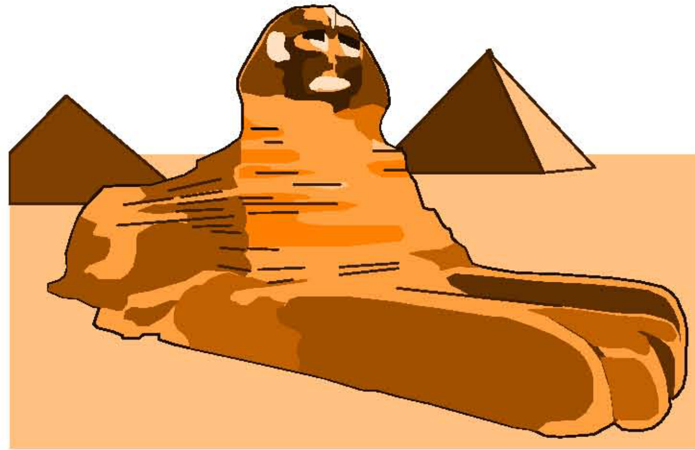
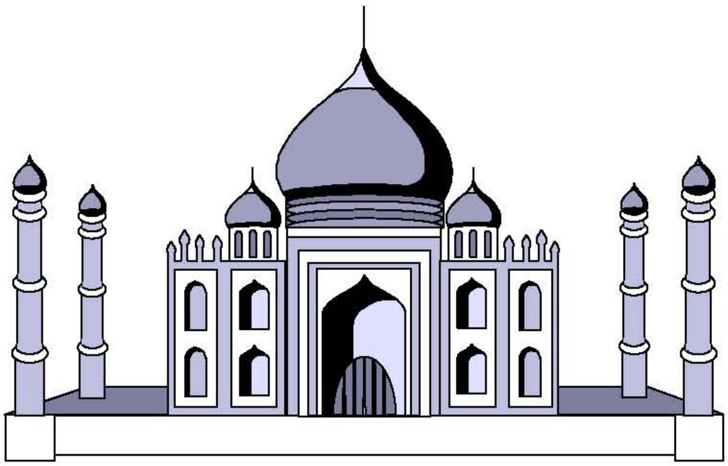
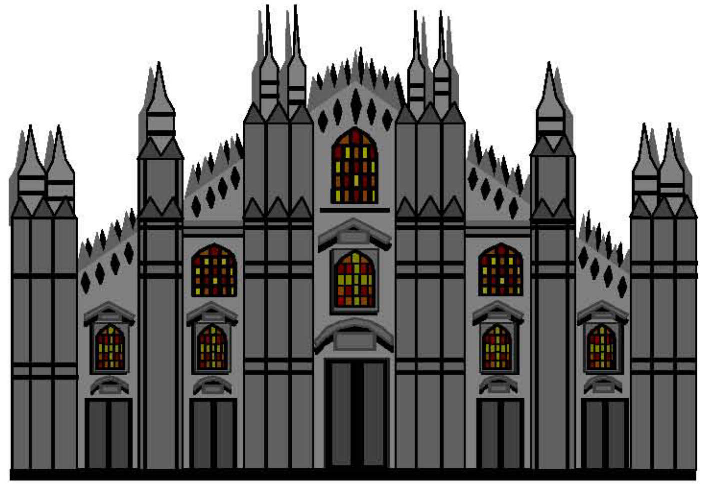


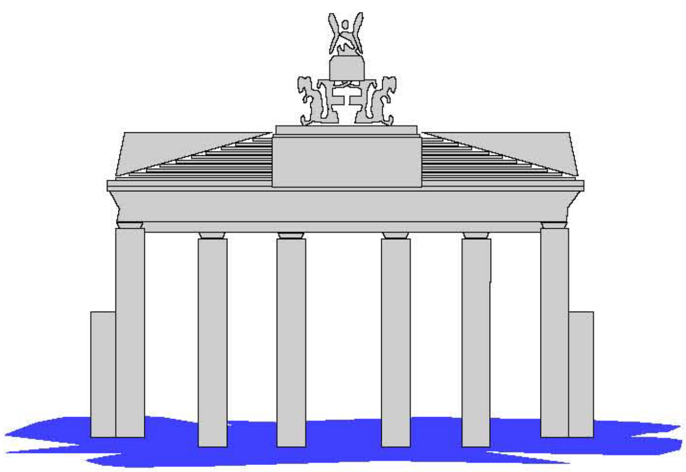
TABLE	HARP
NAIL	TICKET
NEEDLE	PAJAMAS
TOMB	TABLOID

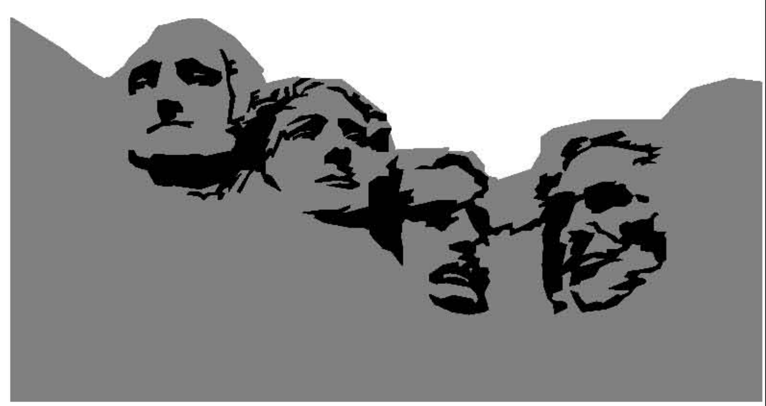






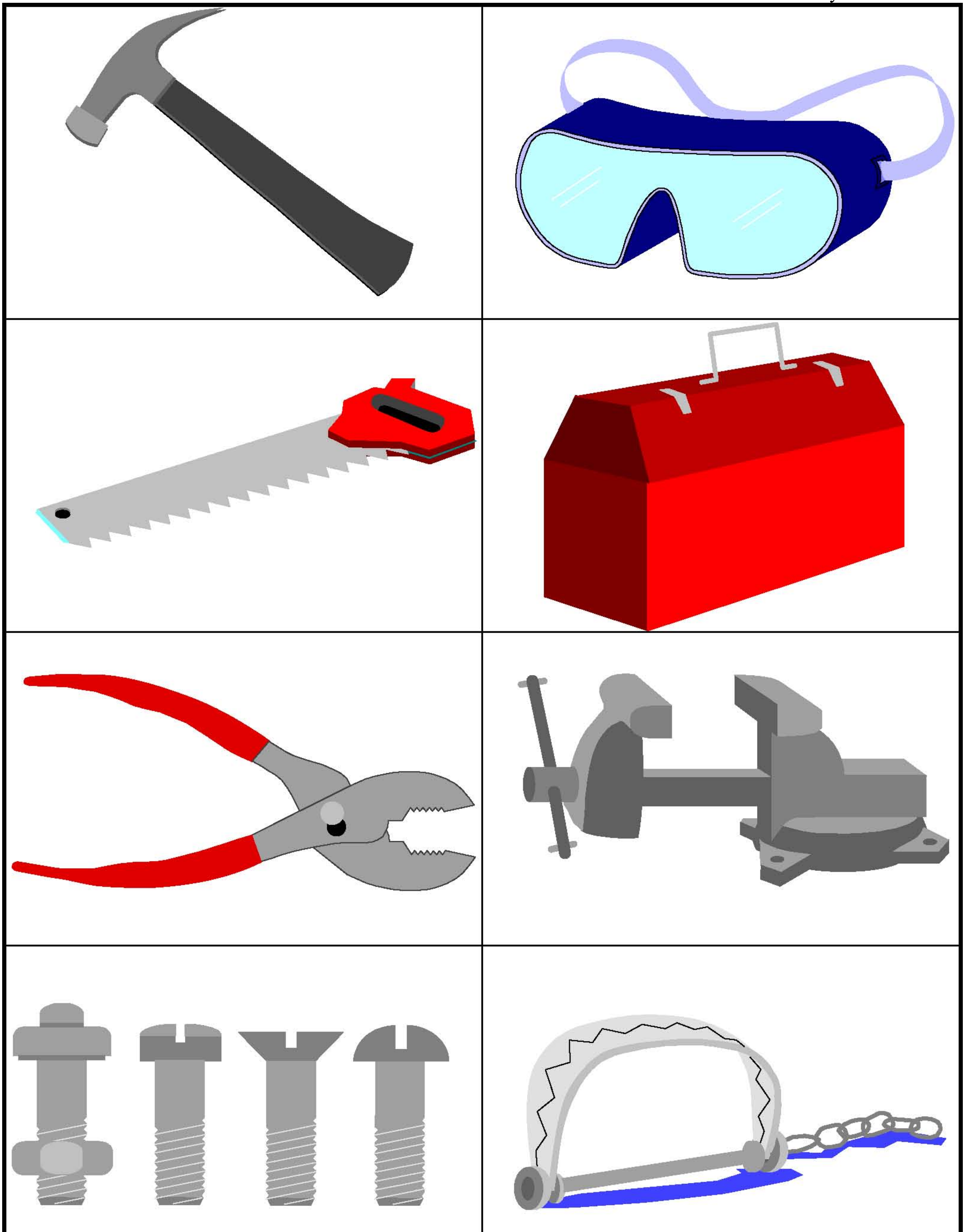




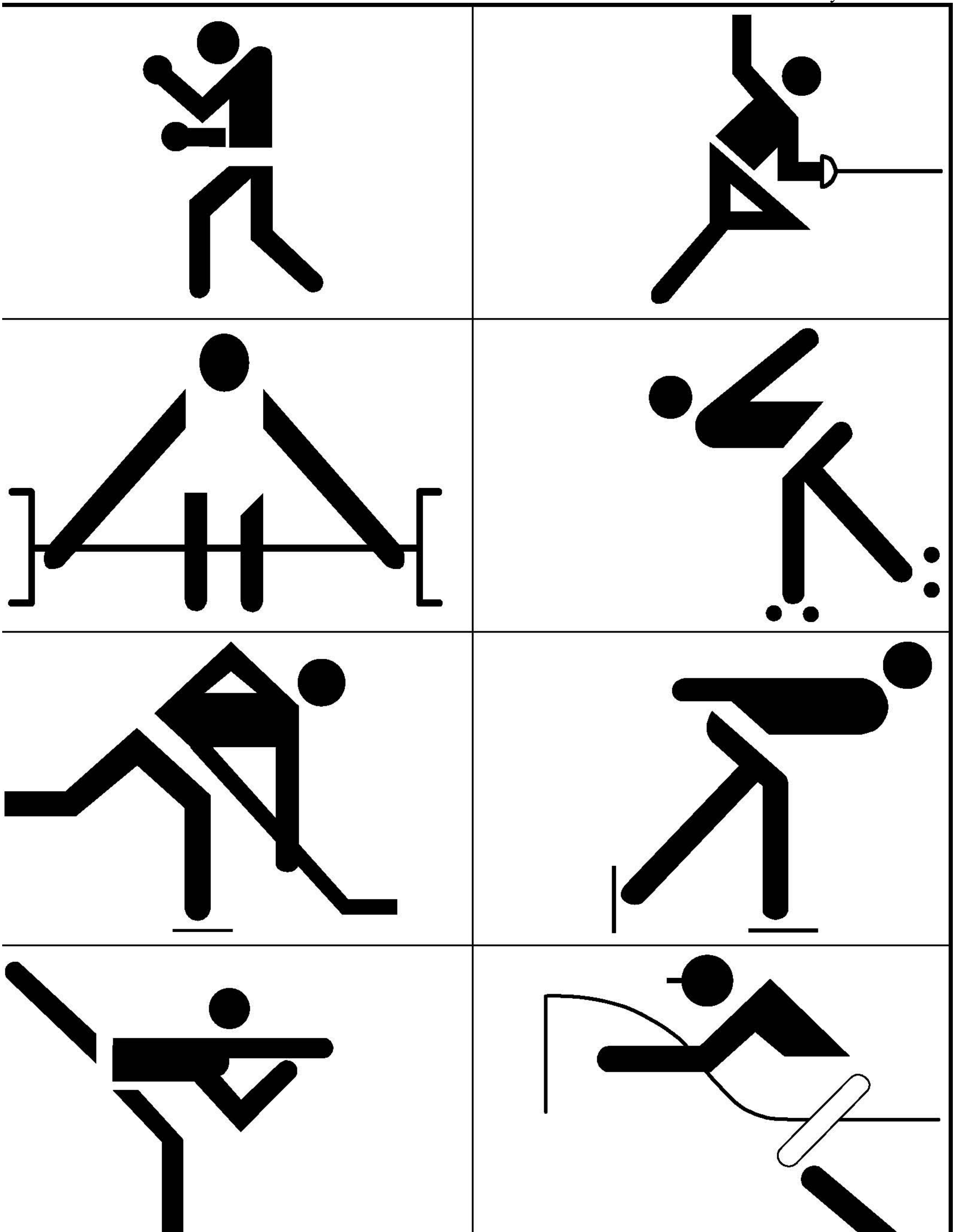


Humpty Dumpty sat on a wall	Jack Sprat could eat no fat
Mary had a little lamb	Jack and Jill went up the hill
Jack be nimble, Jack be quick	Mary, Mary, quite contrary
Three blind mice	Hickory dickory dock

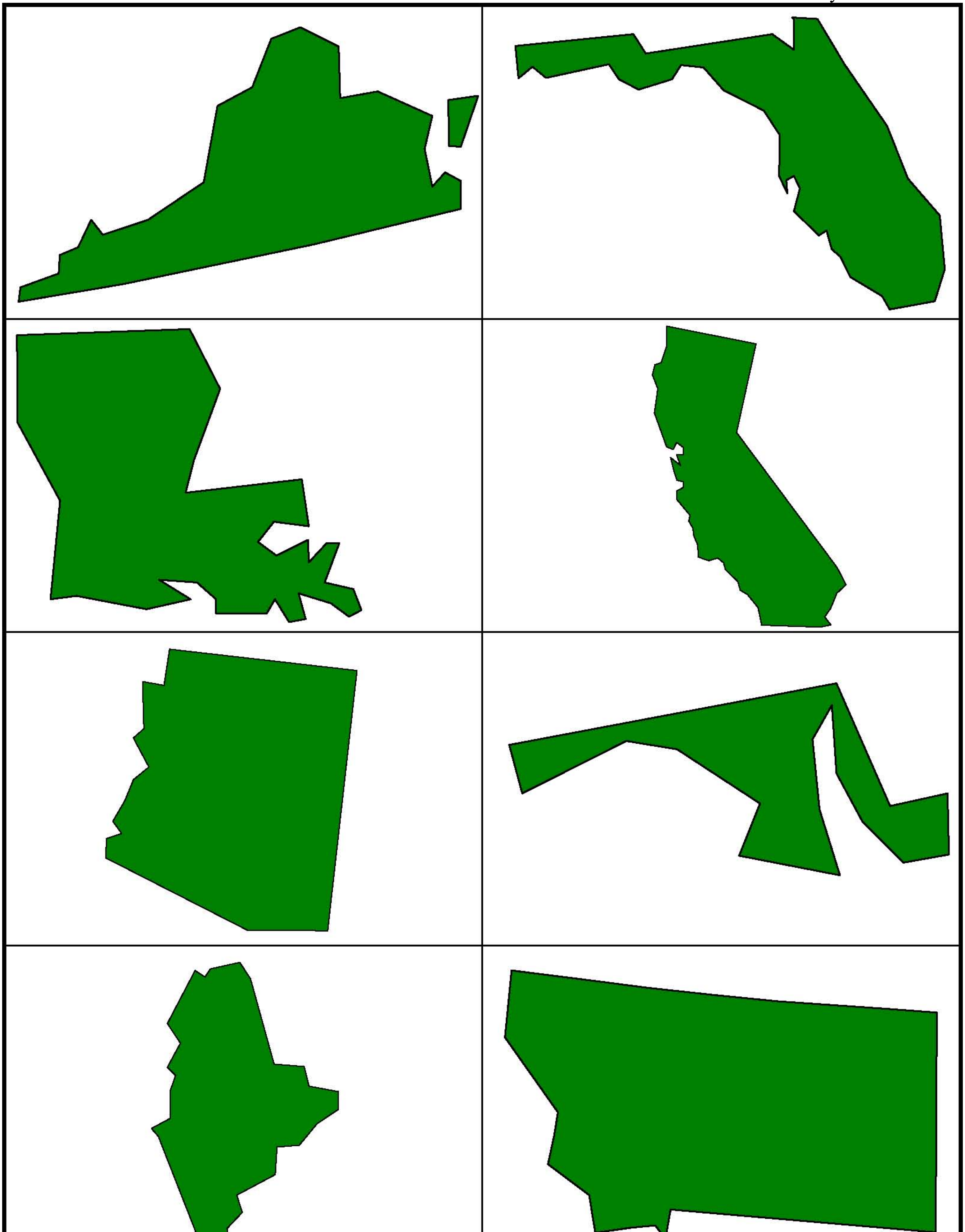
Wada II English Answers



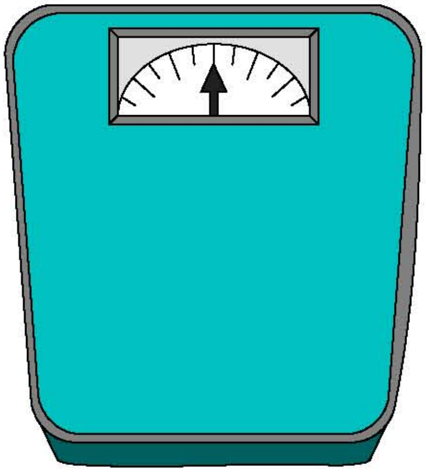
HER	YOURS
SHE	ME
HIM	THESE
ITS	WE

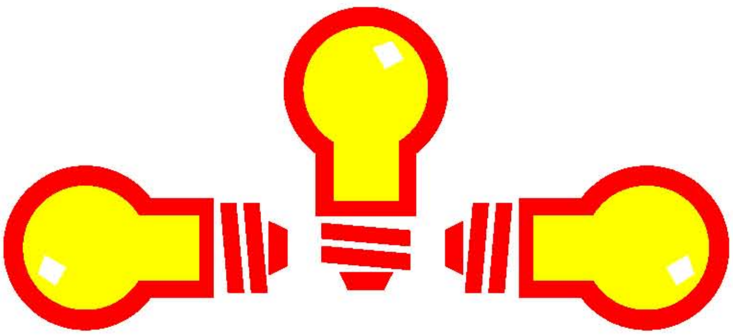


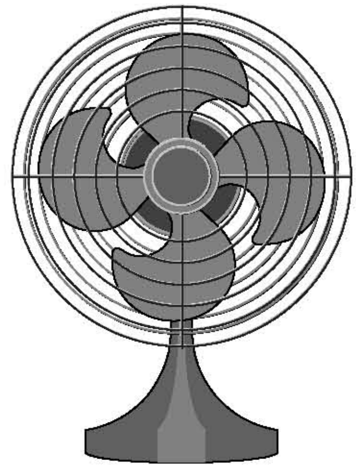
ALLIGATOR	AUTO
APPLE	ASCOT
ARROW	AMPLIFIER
ANGEL	ANTEATER

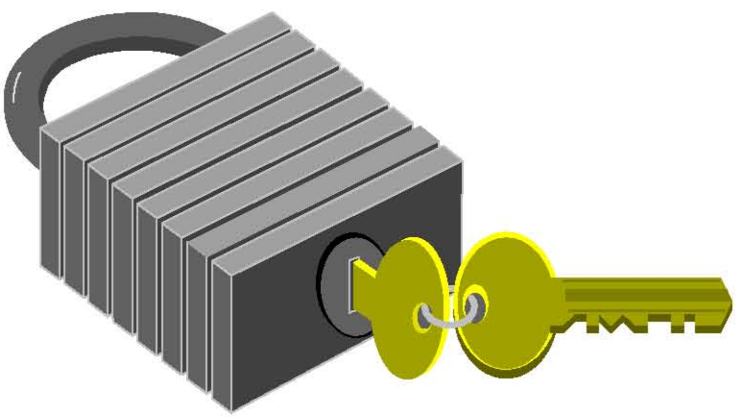


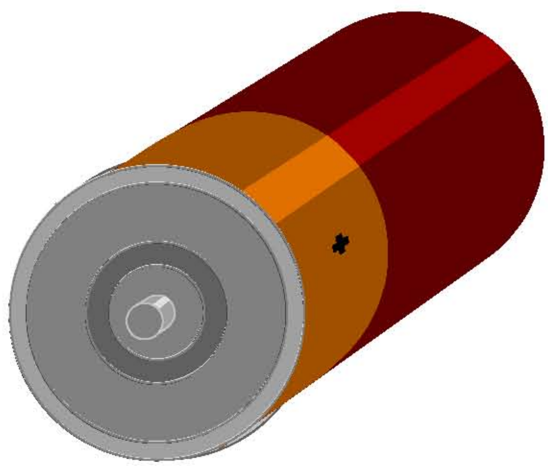
NEED	INDEED
MUCH	MORE
DESTROY	ALONGSIDE
REALLY	FORGONE



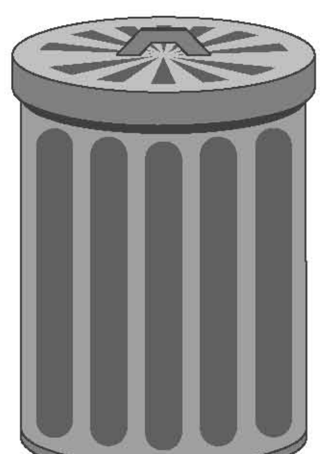




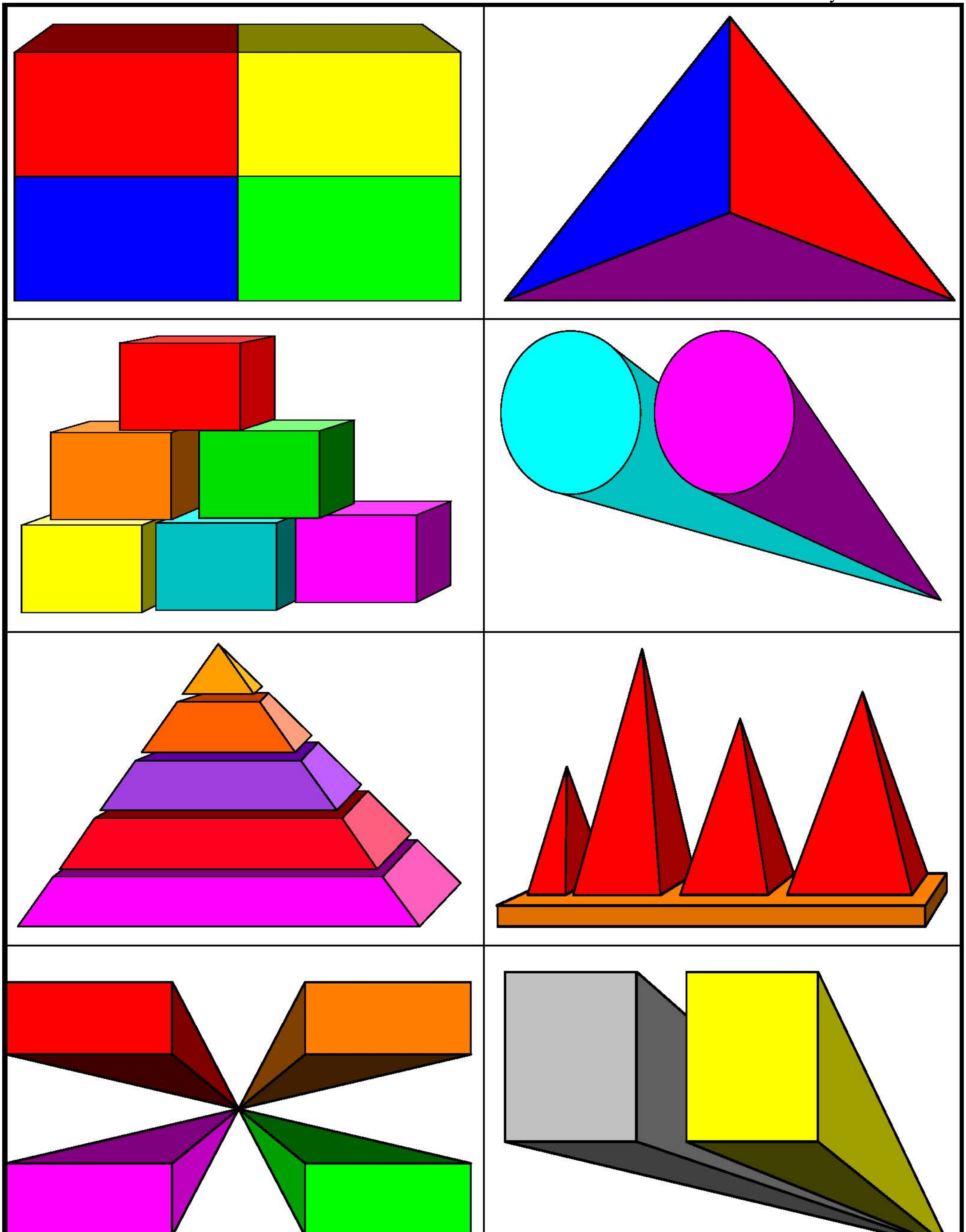




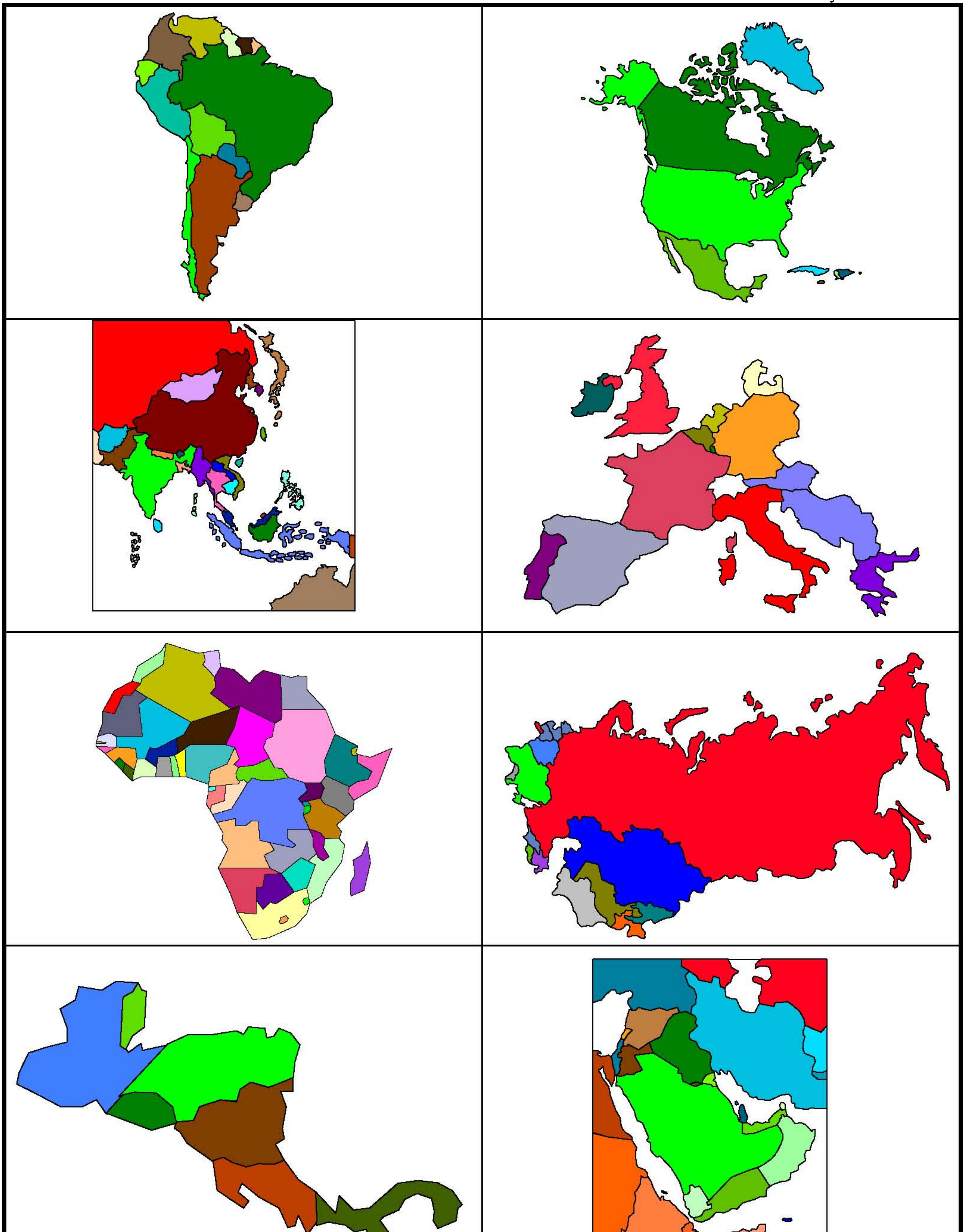

½ CUP



OCTOPUS	HANGER
BED	TREE
PENCIL	DUPLEX
TROPHY	FREIGHTER



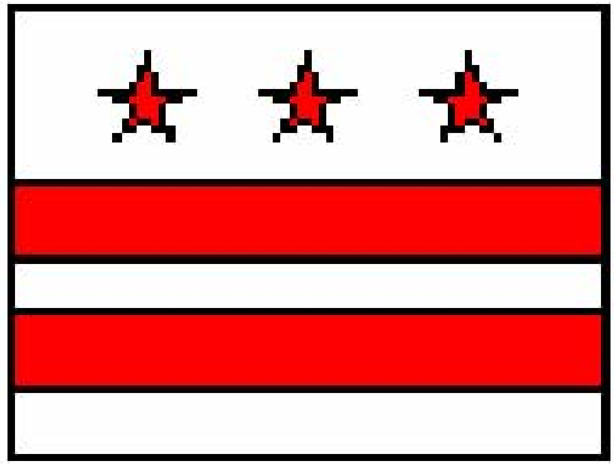
MOREOVER	FORM
DEN	REST
INSIDE	REGARD
BESIDES	EMERGE

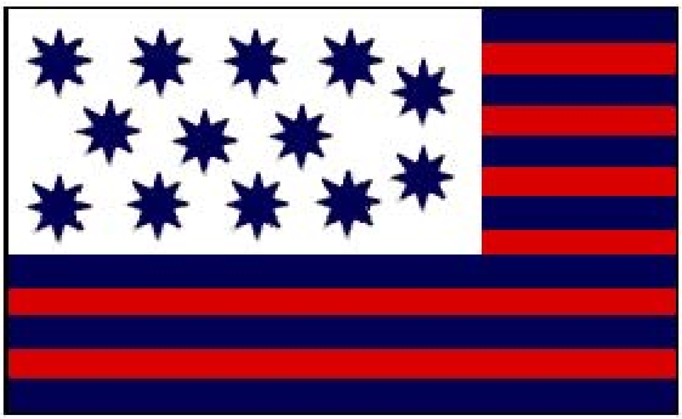


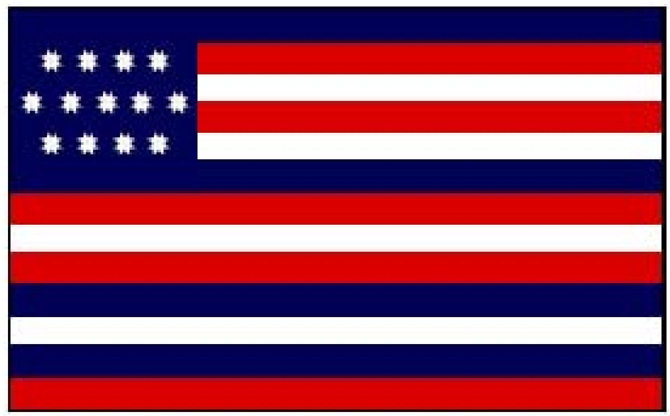
CANOE	WREATH
FEATHER	GLOBE
GALAXY	OBOE
JUPITER	UNICORN

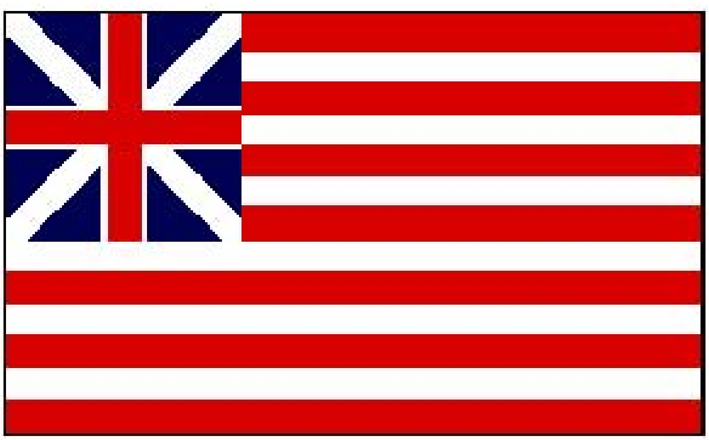
$	Pts
NT$	¥
£	Bs
Rbl	DM

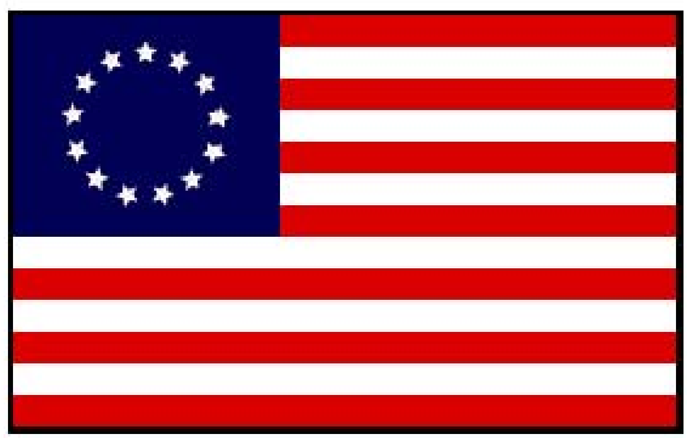
AMONG	THEREIN
JUST	SUCH
DURING	MEAN
BELOW	WHILE

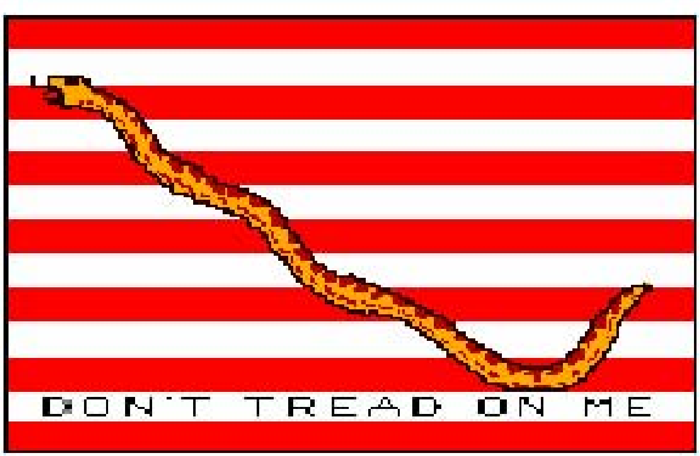






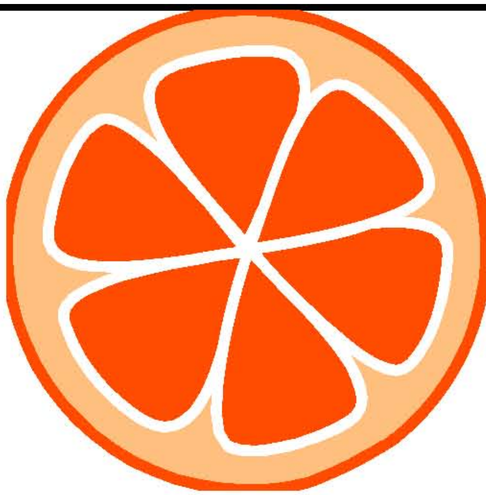
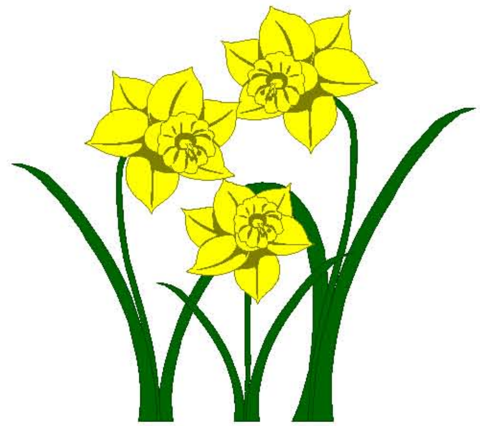


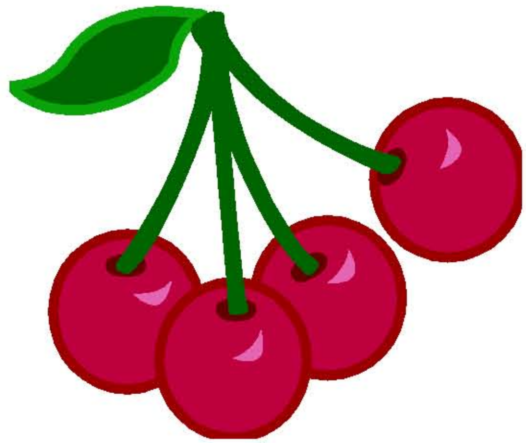
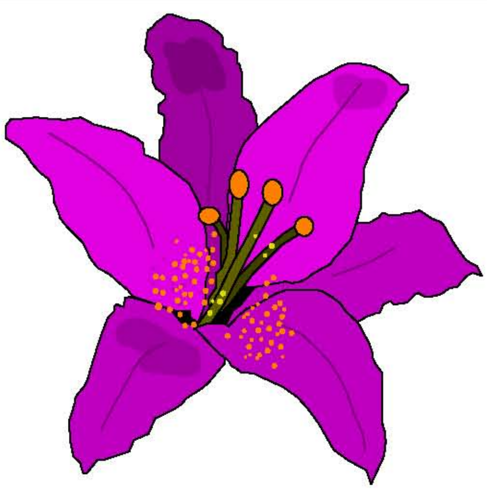
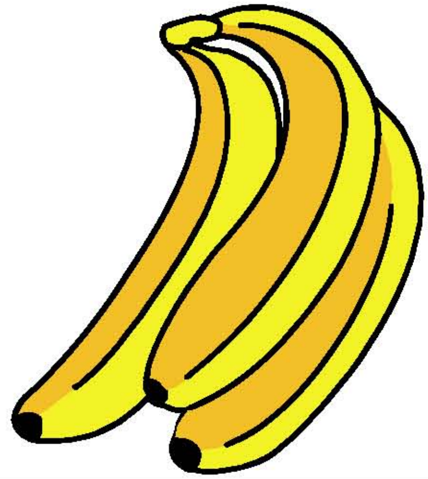
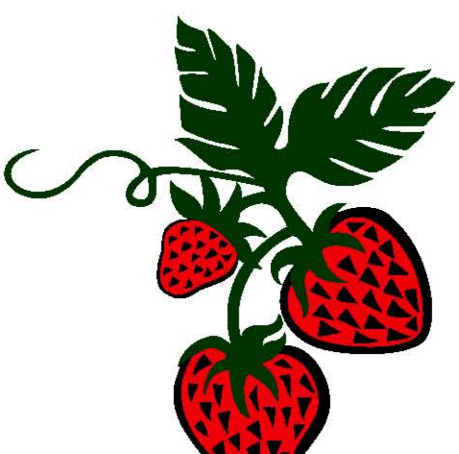



DON'T TREAD ON ME


DON'T TREAD ON ME

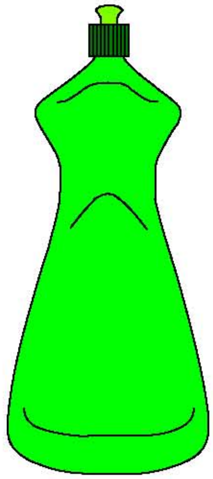
HAMMOCK	PELICAN
SPARROW	MONKEY
MOSQUITO	EAGLE
LOUNGE	ROCKER

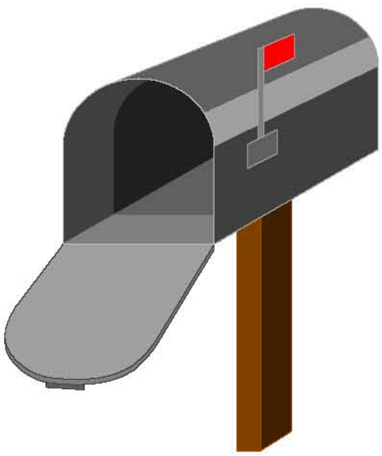



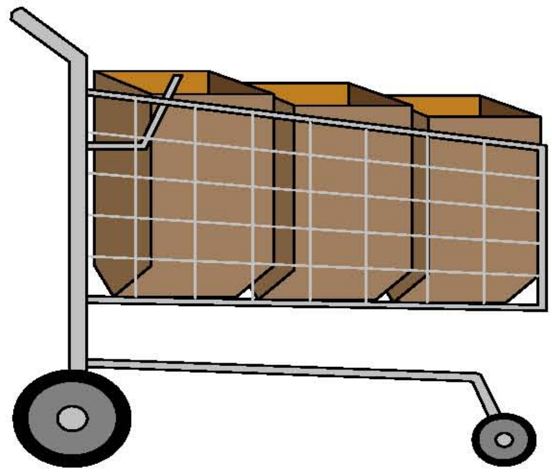





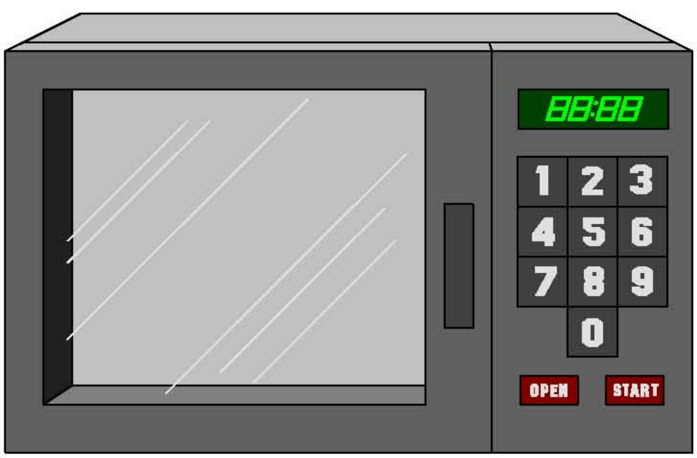

Old MacDonald had a farm	Simple Simon met a pieman
Baa, Baa black sheep	Twinkle twinkle little star
This old man, he played one	Little Tommy Tucker
Quoth the raven, Nevermore	Rain, rain, go away

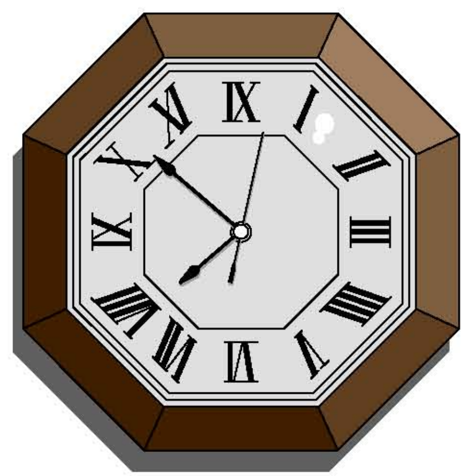
Wada III English Answers

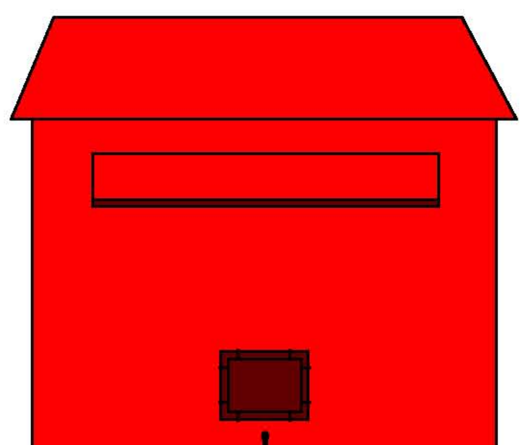




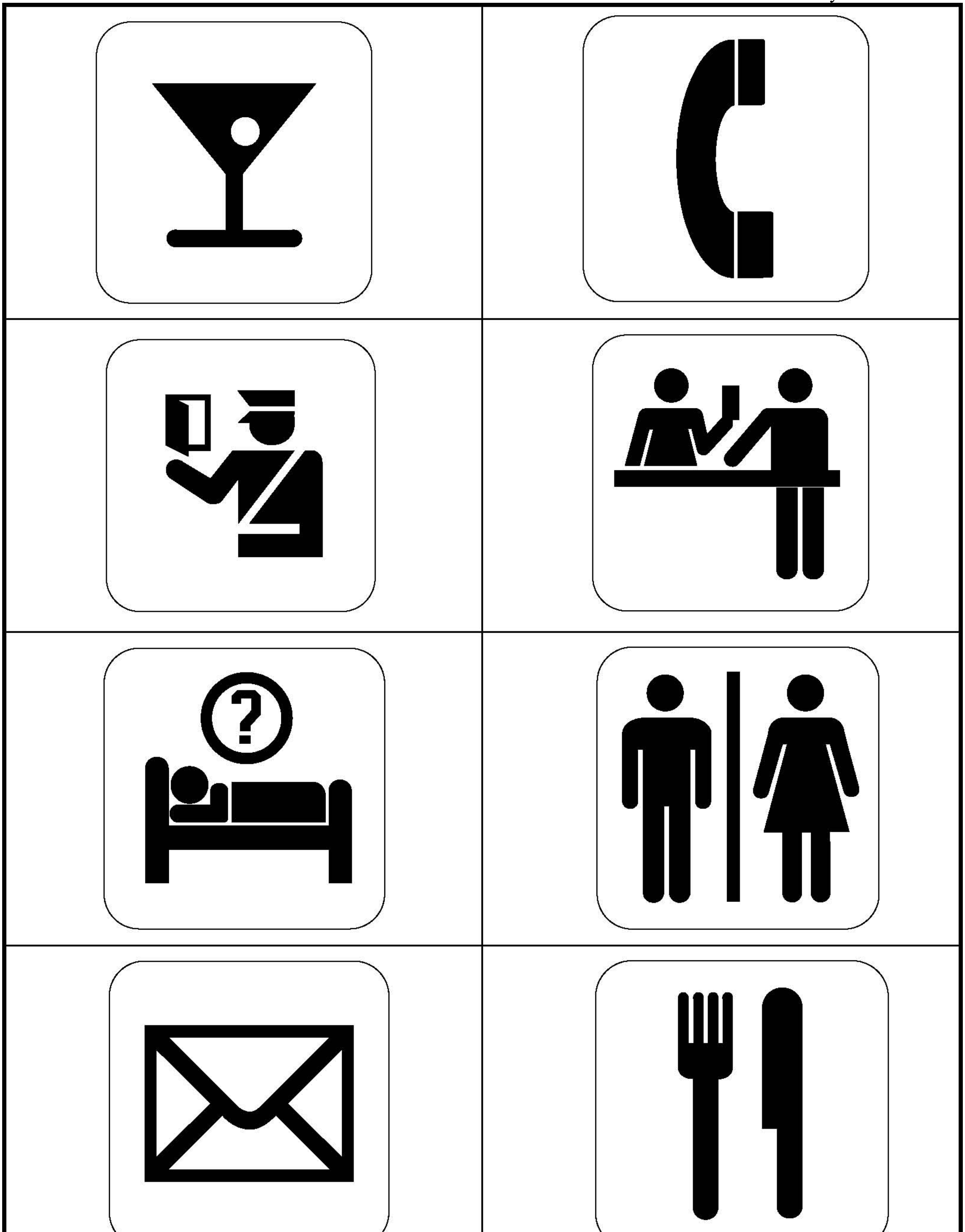



88:88
1 2 3
4 5 6
7 8 9
0
OPEN
START

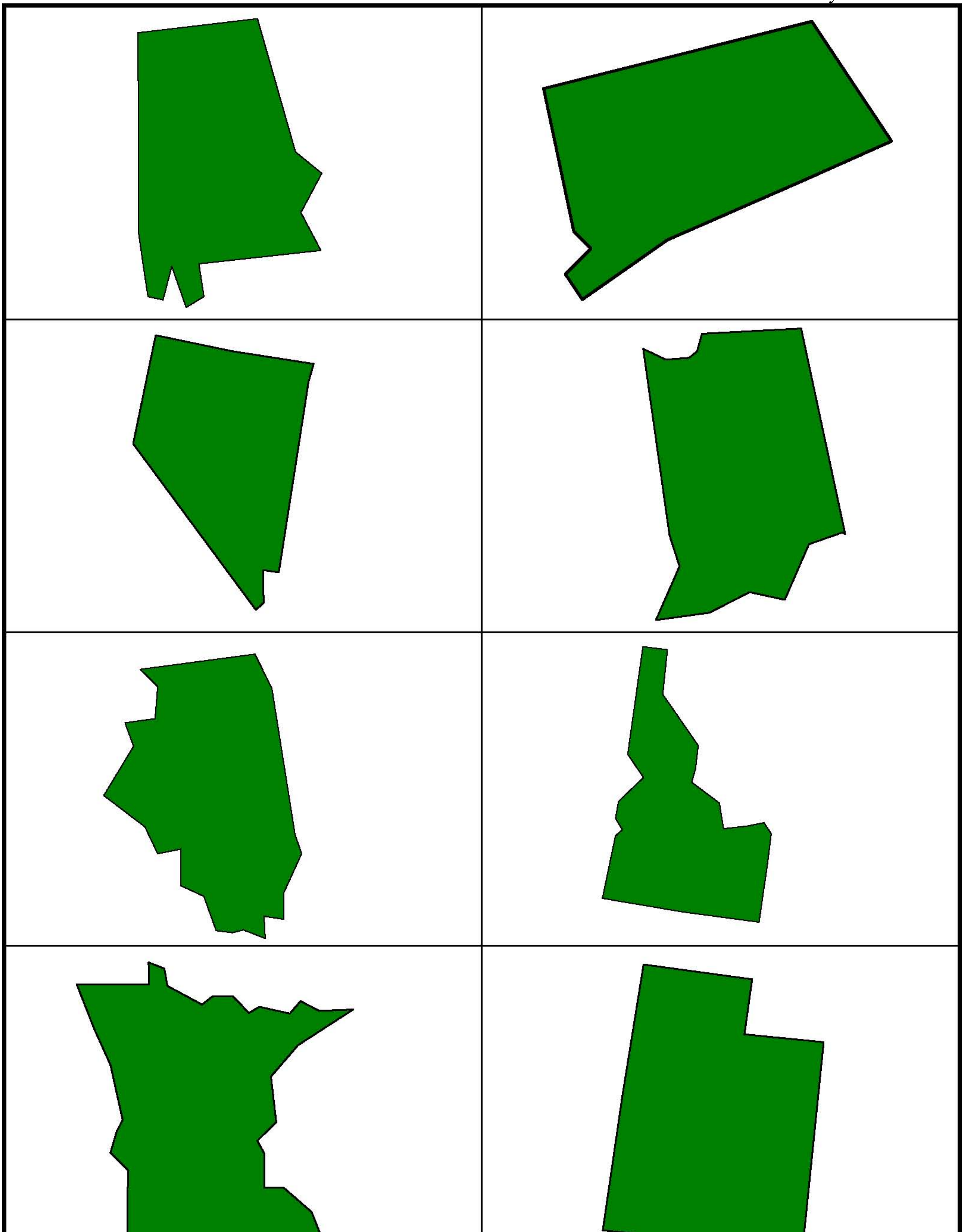




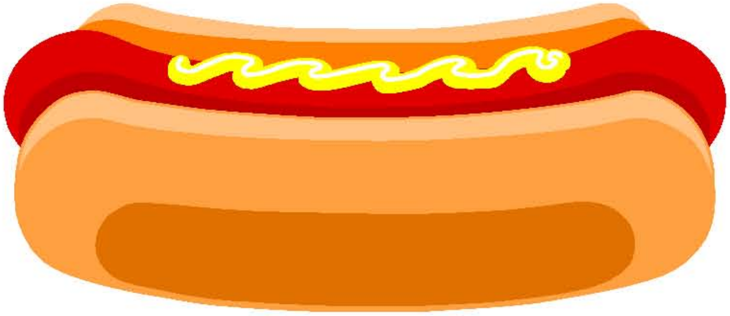
MYSELF	YOU
THEM	HIS
THEIRS	IT
HERSELF	OURS

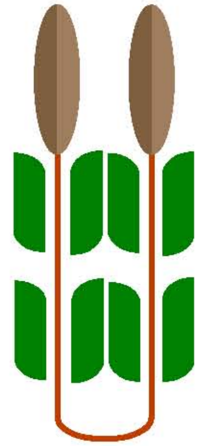


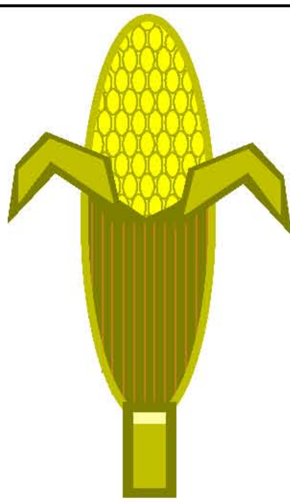
CAT	CHAIR
CAR	COIN
COLUMN	COWBOY
CAMPER	CIGAR

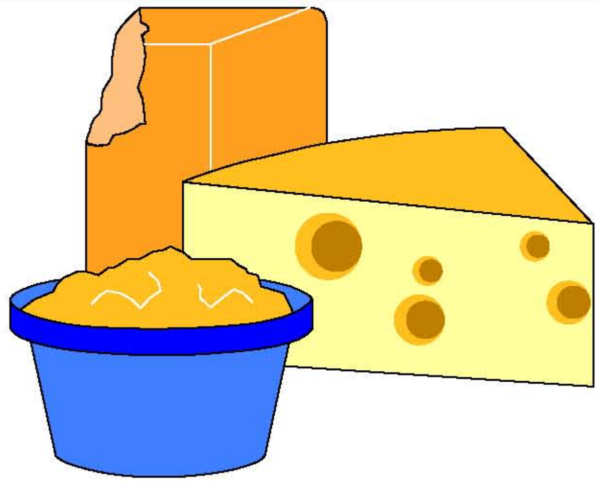


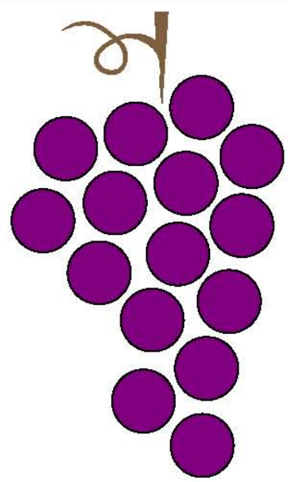
JUSTIFY	BEHIND
AROUND	LATTER
UNLESS	BESIDE
AFTER	OVER

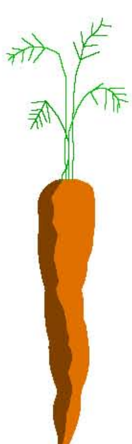


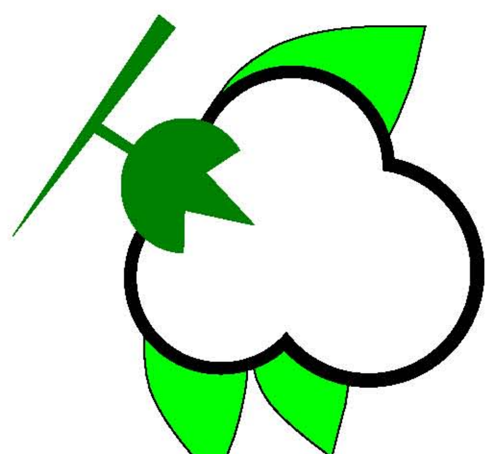




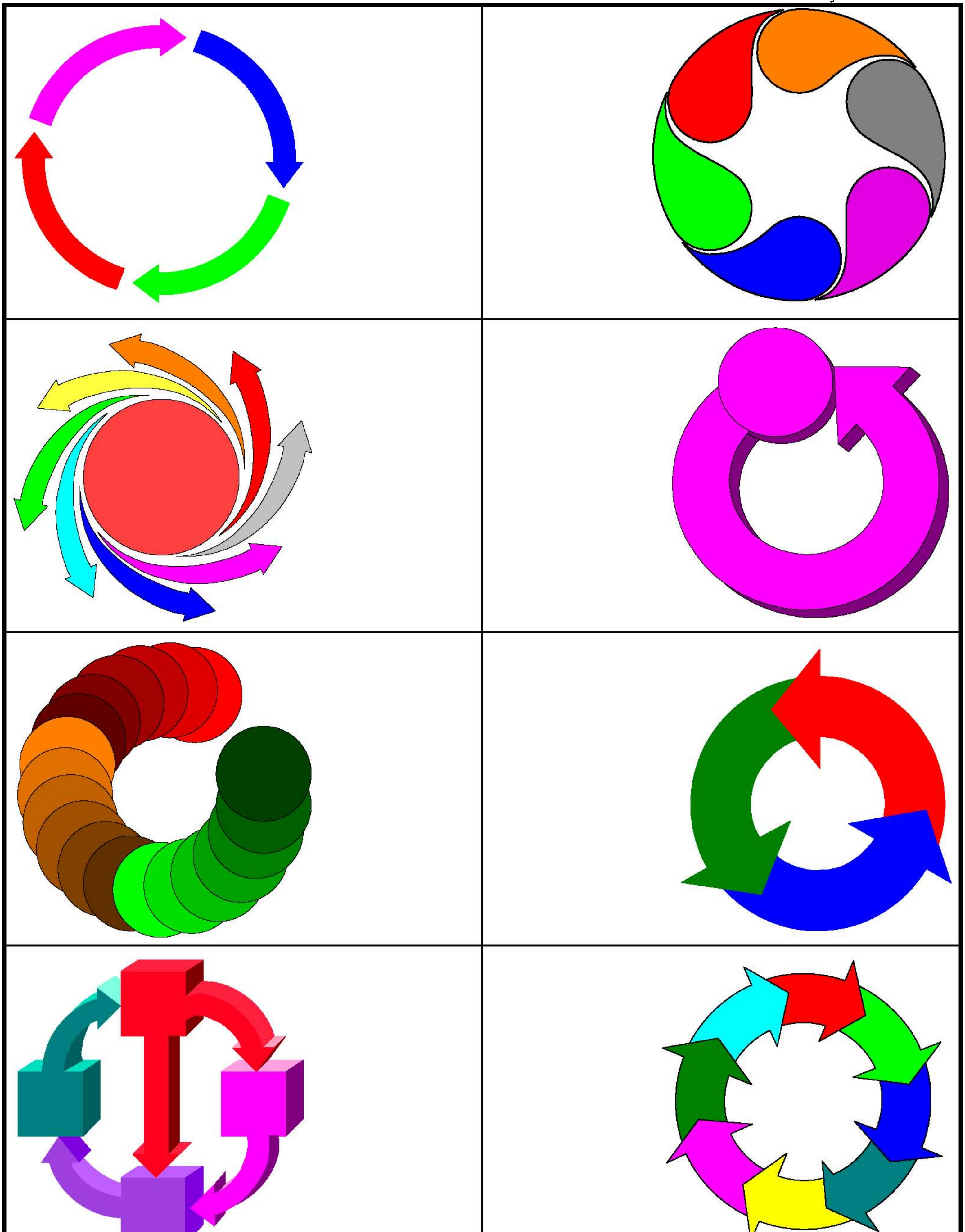




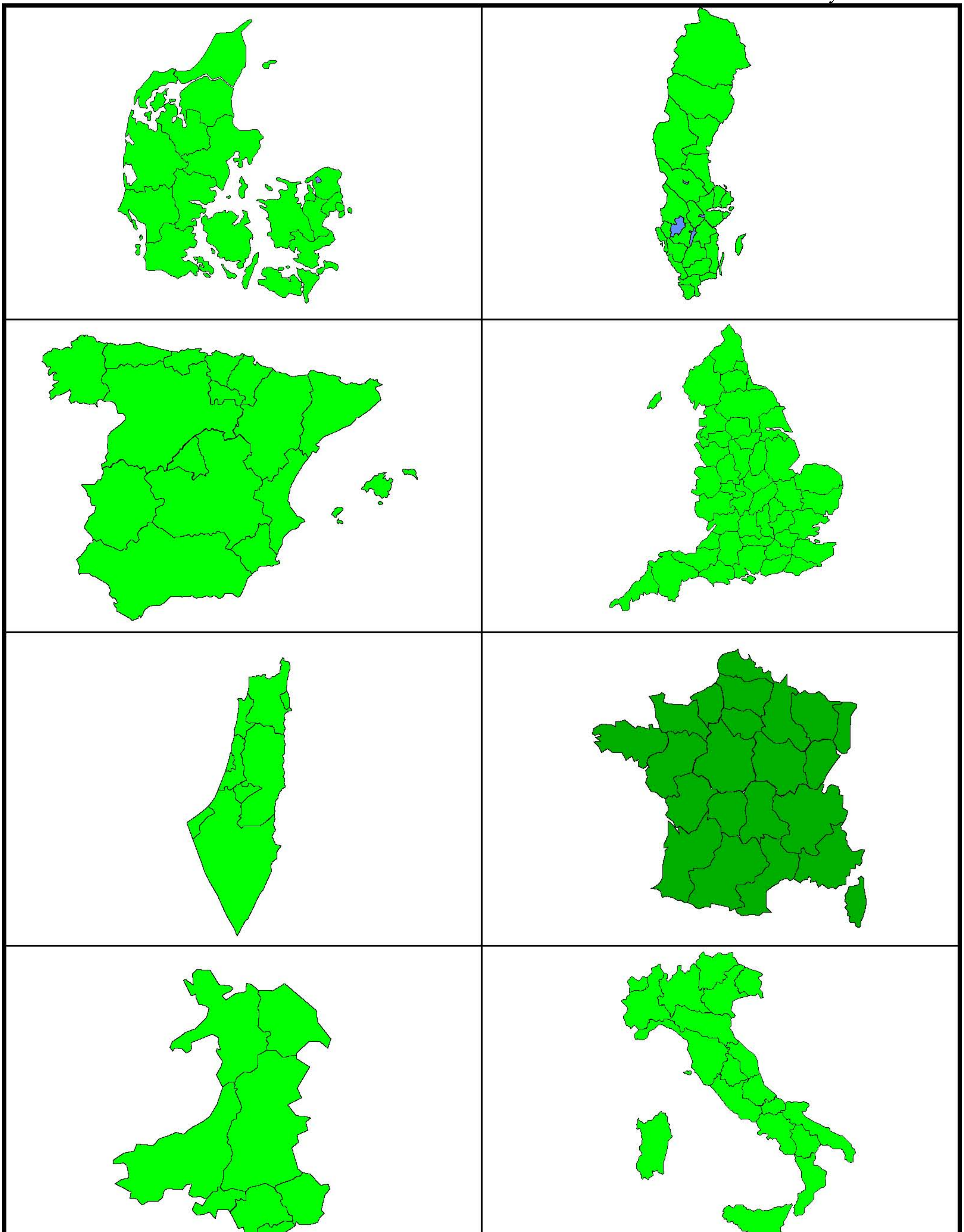




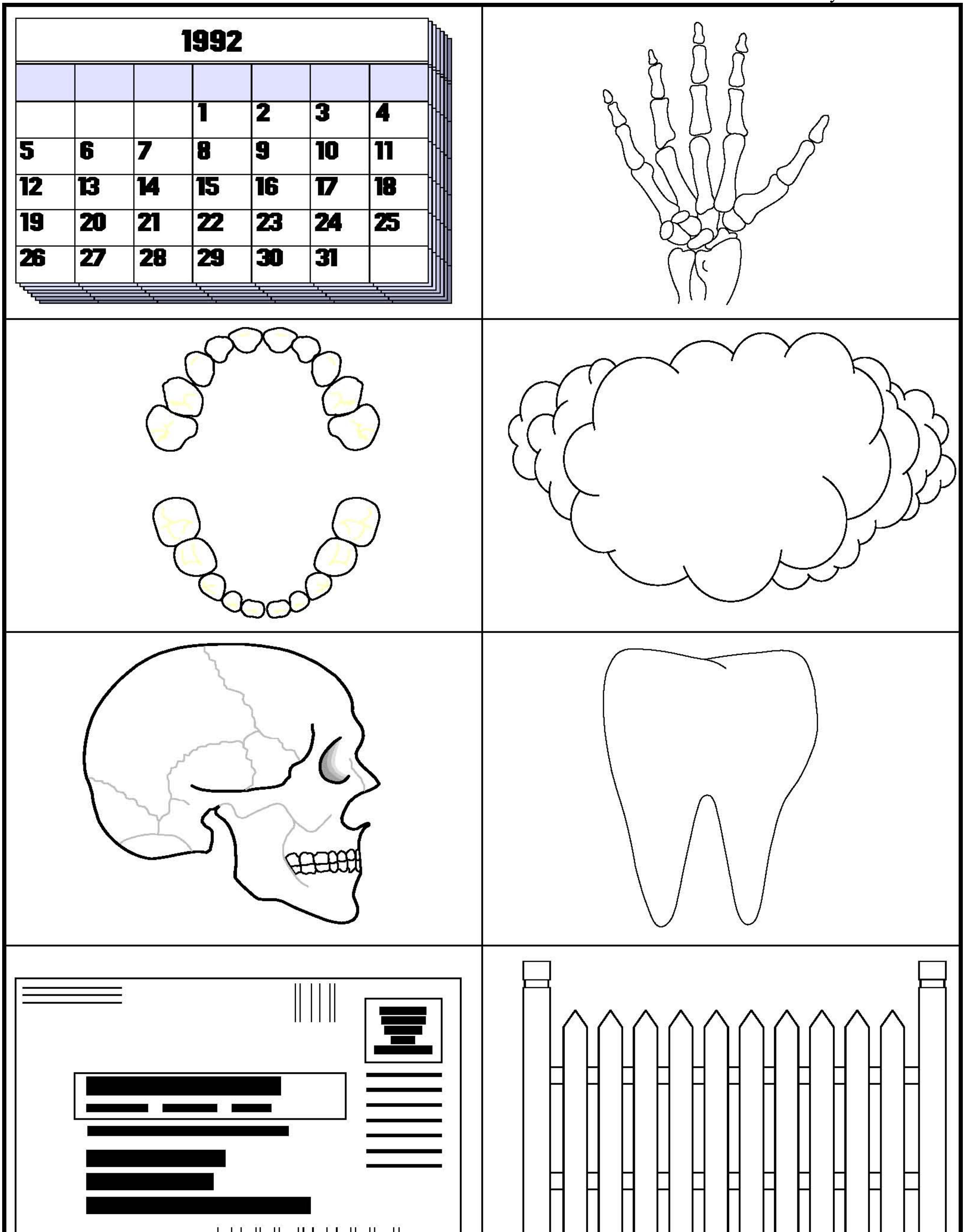
PHONE	WHISTLE
LATCH	COMB
AIRPLANE	COMPASS
ROPE	DAISY



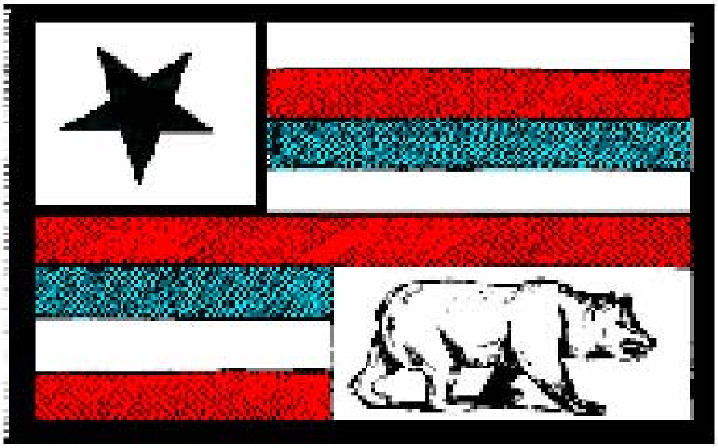
USUAL	GATHER
ABOUT	FURTHER
ALIKE	FORMING
REST	RETRIEVE

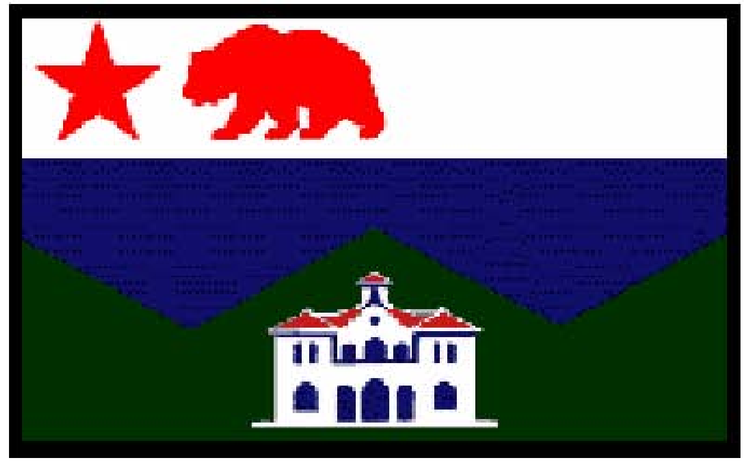


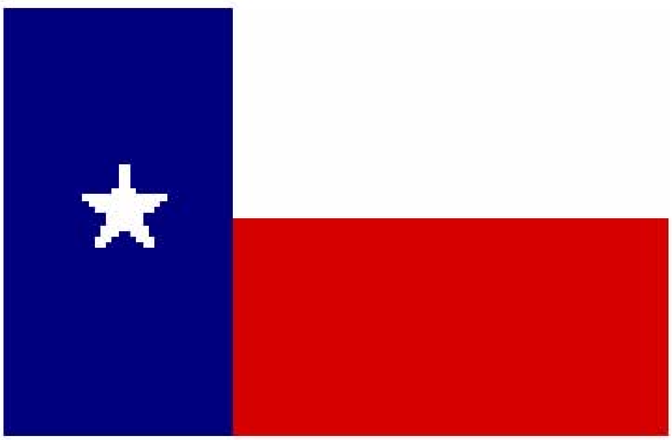
PEPPER	FORK
FROG	PIANO
TOAST	FERN
POTATO	PHOTO

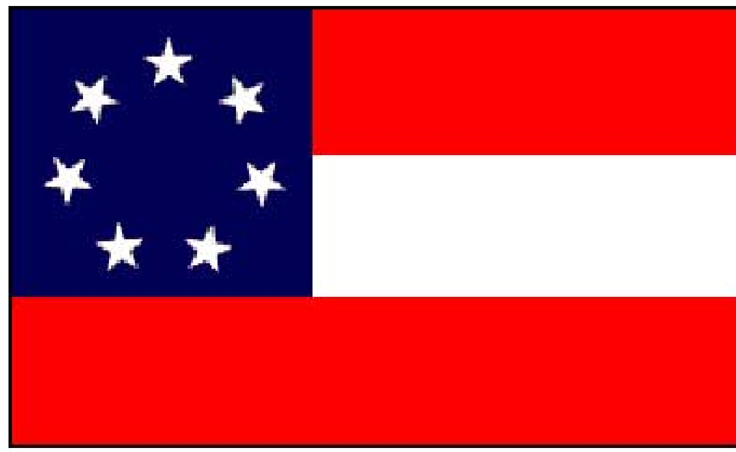

1992
1 2 3 4
5 6 7 8 9 10 11
12 13 14 15 16 17 18
19 20 21 22 23 24 25
26 27 28 29 30 31

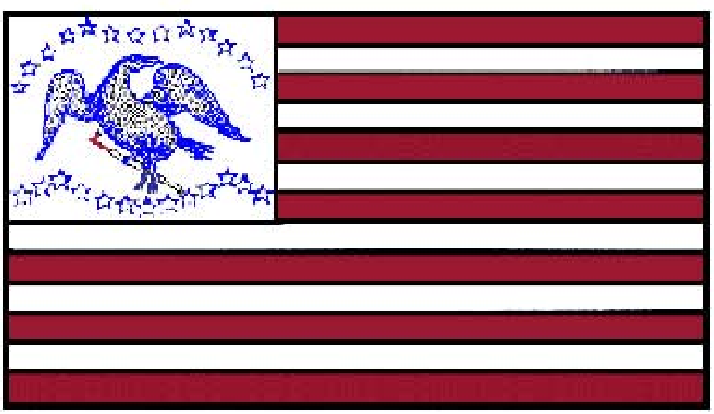
HOLLOW	CREATE
RESTORE	WASH
POINT	UPON
RULE	COMPRESS

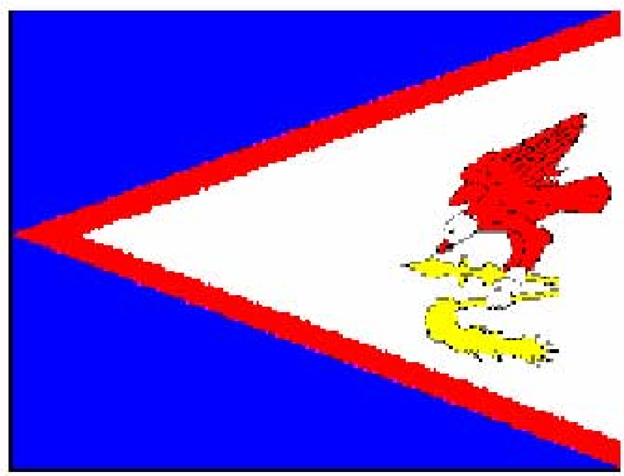




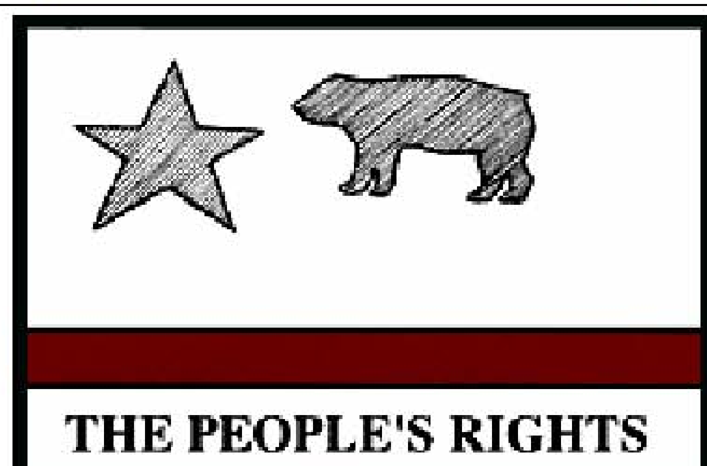




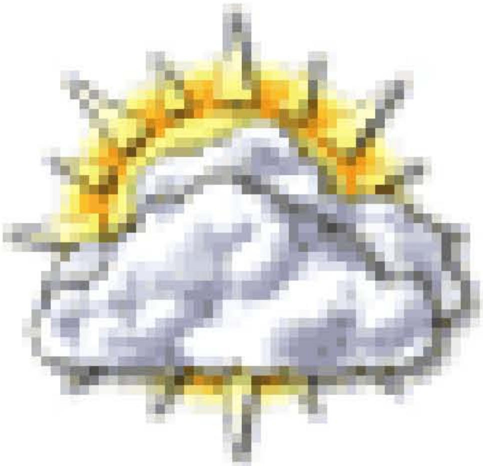
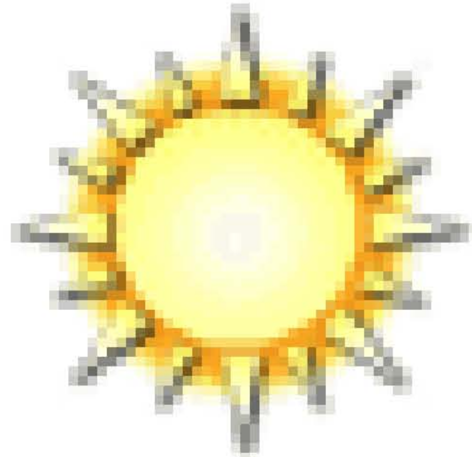
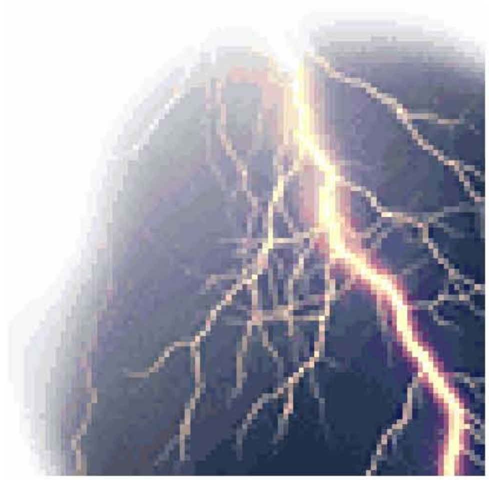
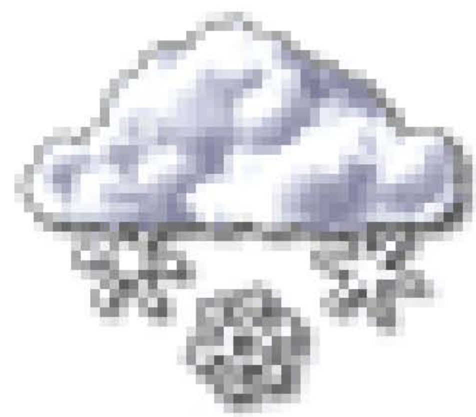





E PLURIBUS UNUM


THE PEOPLE'S RIGHTS

JOURNAL	MEADOW
BEAVER	VIOLIN
GIRAFFE	HYENA
LIBRARY	MAGAZINE

Peter, Peter, pumpkin eater	Roses are red,violets are blue
One, two buckle my shoe	Ring around the rosy
To be, or not to be	Little Miss Muffit
Four and twenty blackbirds	California, here I come

References

Acharya, J. N., & Dinner, D. S. (1997). Use of the intracarotid amobarbital procedure in the evaluation of memory. *Journal of Clinical Neurophysiology, 14*, 311-325.

Alpherts, W. C., Vermeulen, J., & Van Veelen, C. W. (2000). The wada test: prediction of focus lateralization by asymmetric and symmetric recall. *Epilepsy Research, 39*, 239-249.

Anastasi, A. (1988). *Psychological Testing* (6th edition). New York: Macmillian.

Axelrod, B.N., Goldman, B.S., & Woodard, J.L. (1992). Interrater reliability in scoring the Wisconsin Card Sorting Test. *The Clinical Neuropsychologist, 6*, 143-155.

Babb, T.L., & Brown, W.J. (1987) Pathological findings in epilepsy. In J. Jerome Engel (Ed.), *Surgical Treatment of the Epilepsies* (pp. 511-540). New York: Raven Press.

Barr, W. B., Chelune, G.J., Hermann, B.P., Loring, D.W., Perrine, K., Strauss, E., Trenerry, M.R., & Westerveld, M. (1997). The use of figural reproduction tests as measures of nonverbal memory in epilepsy surgery candidates. *Journal of the International Neuropsychological Society*, *3*, 435-443.

Berry, D.T., Allen, R.S., & Schmitt F.A. (1991). The Rey-Osterrieth complex figure: Psychometric characteristic in a geriatric sample. *The Clinical Neuropsychologist, 5*, 143-153.

Branch, C., Milner, B., & Rasmussen, T. (1964). Intracarotid sodium amytal for the lateralization of cerebral speech dominance. *Journal of Neurosurgery, 21*, 399-405.

Breier, J., Plenger, P., Castillo, R., Fuchs, K., Wheless, J., Thomas, A., Brookshire, B., Willmore, L. J., and Papanicolaou, A. (1996). Effects of temporal lobe epilepsy on spatial and figural aspects of memory for a complex geometric figure. *Journal of the International Neuropsychological Society*, *2*, 535-540.

Chang, B. S., & Lowenstein, D. H. (2003). Epilepsy. *New England Journal of Medicine, 349*, 1257-1266.

Chelune, G.J., Naugle, R., Luders, H., & Iwad, I.A. (1991). Prediction of cognitive change as a function of preoperative ability status among temporal lobectomy patients. *Neurology*, *41*, 399-404.

Chelune, G., Naugle, R., Luders, H., Sedlak, J., & Awad, I. A. (1993). Individual change after epilepsy surgery: Practice effects and base-rate information. *Neuropsychology,* 7, 41-52.

Chelune, G. J. (1995). Hippocampal adequacy versus functional reserve: Predicting memory functions following temporal lobectomy. *Archives of Clinical Neuropsychology, 10*, 413-432.

Christianson, S. A., Saisa, J., & Silfvenius, H. (1990). Hemisphere memory differences in Sodium Amytal testing of epileptic individuals. *Journal of Clinical and Experimental Neuropsychology, 12*, 681-694.

Corkin, S. (1984). Lasting consequences of bilateral medial temporal lobectomy: Clinical course and experimental findings in H. M. Kosslyn, Stephen M (Ed); Andersen, Richard A (Ed), (1992). *Frontiers in cognitive neuroscience.* (pp. 516-526). *Cambridge*, US: The MIT Press.

Crocker, L., & Algina, J. (1986). *Introduction to Classical and Modern Test Theory*. Belmont, CA: Wadsworth Group.

Courtney, S.M., Petit, L., Maisog, J.M., Ungerleider, L.G., & Haxby, J.V. (1998). An area specialized for spatial working memory in human frontal cortex. *Science*, *279*, 1347-1351.

Delaney, R.C., Rosen, A.J., Mattson, R.H., & Novelly, R. A. (1980). Memory Function in focal epilepsy: A comparison of non-surgical, unilateral temporal lobe and frontal samples. *Cortex*, *16*, 103-117.

Davenport, J. (1998). Epilepsy. In J. H. Stein (Ed.), *Internal Medicine,* Stein (5 ed.): Mosby, Inc.

Delis, D.C., Kramer, J.H., Kaplan, E., & Ober, B.A. (1987). *California Verbal Learning Test: Research Edition Manual*. San Antonio, TX: The Psychological Corporation, Harcourt Brace and Jovanovich, Inc.

Delis, D.C., Freeland, J., Kramer, J.H., & Kaplan, E. (1988). Integrating clinical assessment with cognitive neuroscience: Construct validation of the California Verbal Learning Test. *Journal of Consulting and Clinical Psychology*, *56*, 123-130.

Devinsky, O. (2004). Diagnosis and Treatment of Temporal Lobe Epilepsy. *Reviews in Neurological Diseases, 1*, 2-9.

Devous, M. D. (1995). SPECT functional brain imaging. Technical considerations. *Journal of Neuroimaging, 5,* Supplemental 1, S2-13.

Dinner, D. S., Luders, H., Morris, H. H., Wyllie, E., & Kramer, R. E. (1987). Validity of intracarotid sodium amobarbital (Wada test) for evaluation of memory function. *Neurology, 37*, 142.

Dodrill, C. B., & Ojemann, G. A. (1997). An exploratory comparison of three methods of memory assessment with the intracarotid amobarbital procedure. *Brain Cognition, 33*, 210-223.

Duchowny, M. S., Harvey, A. S., Sperling, M. R., & Williamson, P. D. (1998). Indications and Criteria for Surgical Intervention. In J. Jerome Engel & T. A. Pedley (Eds.), *Epilepsy a Comprehensive Textbook* (Vol. 2, pp. 1689-1690). Philadelphia, PA: Lippincott-Raven.

Eadie, K., & Shum, D. (1995). Assessment of Visual Memory: A Comparison of Chinese Characters and Geometric Figures as Stimulus Materials. *Journal of Clinical and Experimental Neuropsychology, 17*, 731-739.

Ebel, R. L. (1965). *Measuring educational achievement.* Englewood Cliffs, N.J:. Prentice-Hall.

Engel, J., Jr. (1996). Surgery for seizures. *New England Journal of Medicine, 334*, 647-652.

Engel, J., Jr., & Shewmon, D. A. (1993). *Surgical Treatment of the Epilepsies* (Second ed.). New York: Raven Press.

Epker, M.O., Lacritz, L.H., Ringe, W., Van Ness, P., Agostini, M. & Cullum, C.M. (2000). Wada difference a Wada makes: Comparison of two Wada memory tests. Presented at the annual meeting of the American Epilepsy Society, December 1-6, 2000.

Glosser, G., Saykin, A., Deutsch, G., O'Connor, M., and Sperling, M. (1995). Neural organization of material-specific memory functions in temporal lobe epilepsy patients as assessed by the intracarotid amobarbital test. *Neuropsychology*, *9*, 449-456.

Glowinski, H. (1973). Cognitive deficits in temporal lobe epilepsy: An investigation of memory functioning. *Journal of Nervous and Mental Disease*, *157*, 129-137.

Grote, C.L., Wierenga, BA., Smith, M.C., A.M., Kanner, D.C., Bergen, Geremia, G., & Greenlee, W. (1999). Wada difference a day makes. *Neurology*, *52*, 1577-1582.

Guberman, A. (1994). *An Introduction to Clinical Neurology*. Boston/New

York/Toronto/London: Little, Brown and Company.

Guberman, A., & Bruni, J. (1999). *Essentials of Clinical Epilepsy* (Second ed.). Woburn, MA: Butterworth-Heinemann.

Heaton, R. K., Chelune, G. J., Talley, J. L., Kay, G. G., & Curtiss, G. (1981). *Wisconsin Card Sorting Manual.* Odessa FL: Psychological Assessment Resources, Inc.

Heilbronner, Robert. (1992). The Search for a "Pure" Visual Memory Test: Pursuit of Perfection? *The Clinical Neuropsychologist*, *6*, 105-112.

Heiman, Gary. (1996). *Basic Statistics for the Behavioral Sciences* (Second Ed.). Boston, MA: Houghton Mifflin Company.

Helmstaedter, C., Pohl, C., and Elger, C. (1995). Relations between verbal and nonverbal memory peformance: Evidence of confounding effects particularly in patients with right temporal lobe epilepsy. *Cortex*, *31*, 345-355.

Hermann, B.P., Wyler, A.R., Bush, A.J., & Tabatabai, F.R. (1992). The utlility of the Warrington Recognition Memory Test for temporal lobe epilespy: Pre and post operative results, *Journal of Epilepsy*, *8*, 139-145.

Ivnik, R.J., Sharbrough, F.W., & Laws, E.R. (1987). Effects of anterior temporal lobectomy on cognitive function. *Journal of Clinical Psychology*, *43*, 128-137.

Jones-Gotman, M. (1992). Neuropsychological techniques in the identification of epileptic foci. *Epilepsy Research Supplemental, 5*, 87-94.

Knight, J., & Kaplan, E. (Eds.). (2003). *The Handbook of Rey-Osterrieth Complex Figure Usage: Clinical and Research Applications.* Lutz, FL: PAR, Inc.

Kasper, D. L. (2005). *Harrison's principles of internal medicine* (16 ed.). New York: McGraw-Hill: Medical Publishing Division.

Kimura, D. (1963). Right temporal-lobe damage. *Archives of Neurology*, *8*, 264-271.

Kneebone, A. C., Chelune, G. J., Dinner, D. S., Naugle, R. I., & Awad, I. S. (1995). Intracarotid amobarbital procedure as a predictor of material-specific memory change after anterior temporal lobectomy. *Epilepsia, 36*, 857-865.

Kolb, B., & Whishaw, I. Q. (1990). *Fundamentals of human neuropsychology* (3rd ed.). New York: W. H. Freeman & Company.

Larrabee, G. J., Kane, R. L., Schuck, J. R., & Francis, D. J. (1985). Construct validity of various memory testing procedures. *Journal of Clinical and Experimental Neuropsychology, 7*, 239-250.

Lee, G. P., Park, Y. D., Hempel, A., Westerveld, M., & Loring, D. W. (2002). Prediction of seizure-onset laterality by using Wada memory asymmetries in pediatric epilepsy surgery candidates. *Epilepsia, 43*, 1049-1055.

Lesser, R. P., Dinner, D. S., Luders, H., & Morris, H. H. (1986). Memory for objects presented soon after intracarotid amobarbital sodium injections in individuals with medically intractable complex partial seizures. *Neurology, 36*, 895-899.

Lezak, M.D. (1995). Neuropsychological Assessment. (3rd ed.). New York: Oxford University Press.

Loring, D. W., Hermann, B. P., Perrine, K., Plenger, P. M., Lee, G. P., & Meador, K. J. (1997). Effect of Wada memory stimulus type in discriminating lateralized temporal lobe impairment. *Epilepsia, 38*, 219-224.

Loring, D. W., Lee, G. P., Meador, K. J., & Flanigin, H. F. (1990). The intracarotid amobarbital procedure as a predictor of memory failure following unilateral temporal lobectomy. *Neurology, 40*, 605-610.

Loring, D., Martin, R., Meador, R., Meador, K., & Lee, G. (1990). Psychometric construction of the Rey-Osterrieth Complex Figure: Methodological considerations and interrater reliability. *Archives of Clinical Neuropsychology, 5*, 1-14.

Loring, D. W., Meador, K. J., Lee, G. P., & King, D. W. (1992). *Amobarbital Effects and Lateralized Brain Function*. New York: Springer-Verlag.

Loring, D. W., Meador, K. J., Lee, G. P., & King, D. W. (1995). Wada memory asymmetries predict verbal memory decline after anterior temporal lobectomy. *Neurology, 45*, 1329-1333.

Loring, D. W., Meador, K. J., Lee, G. P., King, D. W., Gallagher, B. B., Murro, A. M., et al. (1994). Stimulus timing effects on Wada memory testing. *Archives of Neurology, 51*, 806-810.

Loring, D. W., Meador, K. J., Lee, G. P., & Nichols, M. E. (1994). Wada memory performance predicts seizure outcome following anterior temporal lobectomy. *Neurology, 44*, 2322-2324.

Loring, D. W., Murro, A. M., Meador, K. J., Lee, G. P. (1993). Wada memory testing and hippocampal volume measurements in the evaluation for temporal lobectomy. *Neurology, 43*, 1789-1793.

Loring, D. W., Hermann, B. P., Meador, K. J., Lee, G. P., Gallagher, B. B., King, D. W., et al. (1994). Amnesia after unilateral temporal lobectomy: a case report. *Epilepsia, 35*, 757-763.

Loring, D. W., Murro, A. M., Meador, K. J., Lee, G. P., Gratton, C. A., Nichols, M. E., et al. (1993). Wada memory testing and hippocampal volume measurements in the evaluation for temporal lobectomy. *Neurology, 43*, 1789-1793.

Loring, D. W., Meador, K. J., Lee, G. P., King, D. W., Gallagher, B. B., Murro, A. M., et al. (1994). Stimulus timing effects on Wada memory testing. *Archives of Neurology, 51*, 806-810.

Loring, D. W., Meador, K.J., Lee, G.P., Nichols, M.E., King, D. W., Gallagher, B.B., Murro, A.M., and Smith, J.R., (1994). Wada memory performance predicts seizure outcome following anterior temporal lobectomy. *Neurology*, *44*, 2322-2324.

Martin, R. C., & Grote, C. L. (2002). Does the Wada test predict memory decline following epilepsy surgery? *Epilepsy & Behavior, 3*, 4-15.

McGlone, J., & MacDonald, B. H. (1989). Reliability of the sodium amobarbital test for memory. *Journal of Epilepsy,* 31-39.

Miller, L.A., Munoz, D.G., & Finmore, M. (1993). Hippocampal sclerosis and human memory. *Archives of Neurology*, *50*, 391-394.

Milner, B. (1972). Disorders of learning and memory after temporal lobe lesions in man. *Clinical Neurosurgery, 19*, 421-446.

Milner, B., Branch, C., & Rasmussen, T. (1962). Study of Short-Term Memory after Intracarotid Injection of Sodium Amytal. *Transactions of the American Neurological Association, 87*, 224-226.

Mungas, D., Blunden, D., Bennington, K., Stone, A., & Palma, G. (1990). Reliability and validity of scales for assessing behavior in epilepsy. *Psychological Assessment, 2*, 423-431.

Naugle, R., Chelune, G., Cheek, R., Luders., and Awad, I. (1993). Detection of changes in material-specific memory following temporal lobectomy using the Wechsler Memory Scale-Revised. *Archives of Clinical Neurology*, *8*, 381-395.

Novelly, R. A., & Williamson, P. D. (1989). Incidence of false-positive memory impairment in the intracarotid amytal procedure. *Epilepsia, 30*, 711.

Nunnaly, J. (1978). *Psychometric theory*. New York: McGraw-Hill.

Osterrieth, P.A. (1944). Le test de copie d'une figure complex: Contribution a l'edude de la perception et de la memoire. *Archives de Psychologie*, *30*, 286-356.

Perrine, K., Sass, K. J., Choi, I. S., Brown, E. R., Luciano, D. J., & Devinsky, O. (1993). Interhemispheric difference scores in IAP memory testing. *Journal of Clinical and Experimental Neuropsychology, 15*, 24.

Perrine, K., Westerveld, M., Sass, K. J., Devinsky, O., Dogali, M., Spencer, D. D., et al. (1995). Wada memory disparities predict seizure laterality and postoperative seizure control. *Epilepsia, 36*, 851-856.

Rasmussen, T., & Milner, B. (1977). The role of early left-brain injury in determining lateralization of cerebral speech functions. *Annals NY Academy of Sciences, 299*, 355-369.

Rausch, R., Babb, T. L., Engel, J., Jr., & Crandall, P. H. (1989). Memory following intracarotid amobarbital injection contralateral to hippocampal damage. *Archives of Neurology, 46*, 783-788.

Rausch, R., and Babb, T., (1993). Hippocampal neuron loss and memory scores before and after temporal lobe surgery for epilepsy. *Archives of Neurology*, *50*, 812-817.

Rausch, R., Le, M.-T., & Langfitt, J. T. (1997). Neuropsychological Evaluation-Adults. In J. Engel, Jr. (Ed.), *Epilepsy: A Comprehensive Textbook*. Philadelphia: Lippincott-Raven Publishers.

Rausch, R., Silfvenius, H., Wieser, H.-G., Dodrill, C. B., Meador, K. J., & Jones-Gotman, M. (1993). Intraaterial amobarbital procedures. In J. E. Jr. (Ed.), *Surgical treatment of epilepsies* (second ed., pp. 341-357). New York: Raven Press.

Reitan, R. (1955). The relation of the Trail Making Test to organic brain damage. *Journal of Consulting Psychology, 19*, 393-394.

Roman, D. D., Beniak, T. E., & Nugent, S. (1996). Memory performance on the intracarotid amobarbital procedure as a predicator of seizure focus. *Epilepsy Research, 25*, 243-248.

Sadock, B. J., & Sadock, V. A. (2003). *Synopsis of Psychiatry* (Ninth ed.). Philadelphia, PA: Lippincott Williams and Wilkins.

Simkins-Bullock, J. (2000). Beyond speech lateralization: a review of the variability, reliability, and validity of the intracarotid amobarbital procedure and its nonlanguage uses in epilepsy surgery candidates. *Neuropsychology Review, 10*, 41-74.

Smith, G.E., Malec, J.F., Ivnik, R.J. (1992). Validity of the construct of nonverbal memory: A factor-analytic study in a normal elderly sample. *Journal of Clinical and Experimental Neuropsychology*, *14*, 211-221.

Snyder, P. J., & Harris, L. J. (1997). The intracarotid amobarbital procedure: an historical perspective. *Brain Cognition, 33*, 18-32.

Snyder, P. J., & Nussbaum, P. D. (1998). *Clinical Neuropsychology*. Washington, DC: American Psychological Association.

Spencer, S. (2002). When should temporal-lobe epilepsy be treated surgically? *The Lancet Neurology, 1*, 375-382.

Spreen, O., & Strauss, E. (1998). *A compendium of neuropsychological tests. (2nd ed)*. New York: Oxford University Press.

Stanulis, R. G., Valentine, R. V., & DeToledo, D. B. (1990). Intracarotid Sodium Amytal memory testing: Use and misuse of results in surgical decisions. *Epilepsia, 31*, 685.

Taylor, L.B., (1969). Localization of cerebral lesions by psychological testing. *Clinical Neurosurgery*, *16*, 269-287.

Trenerry, M. R. (1996). Neuropsychologic Assessment in Surgical Treatment of Epilepsy. *Mayo Clinical Proceeding, 71*, 1196-1200.

Trenerry, M. R., & Loring, D. W. (1995). Intracarotid amobarbital procedure. *Neuroimaging Clinics of North America, 5*, 721-728.

Wada, J. (1949). A new method for the determination of the side of cerebral speech dominance. A preliminary Report on the intracarotid injeciton of sodium amytal in ma. *Medicine and Biology, 14*, 221-222.

Wechsler, D. (1997). *WAIS-III/WMS-III Technical Manual*. San Antonio, TX: The Psychological Corporation/Harcourt Brace & Company.

Wieser, H.G., Engel J., Jr., Williamson, P.D., Babb, T.L., & Gloor, P. (1993). Surgically Remediable Temporal Lobe Syndromes. In Engel, J., Jr. (Ed), *Surgical Treatment of the Epilepsies: Second Edition* (p. 49-63). New York, NY: Raven Press.

Woods, R. P., Dodrill, C. B., & Ojemann, G. A. (1988). Brain injury, handedness, and speech lateralization in a series of amobarbital studies. *Annuls of Neurology, 23*, 510-518.

Wyllie, E. (2001). *The Epidemiology of Epilepsy* (Third edition). Philadelphia, PA: Lippincott Williams & Wilkins.

Wyllie, E., Naugle, R., Chelune, G., Luders, H., Morris, H., & Skibinski, C. (1991). Intracarotid amobarbital procedure: II. Lateralizing value in evaluation for temporal lobectomy. *Epilepsia, 32*, 865-869.

VITAE

Daniel David Eisenman was born in Denver, Colorado, on October 31, 1971, the son of Rabbi Bernard Eisenman and Billie Eisenman. Daniel moved to Massachusetts in 1986 and attended Lexington High School. After completing high school, he entered Clark University in Worcester, Massachusetts. He received a degree of Bachelor of Arts with a major in English Literature. In May 1999 he obtained a Master of Science in Clinical Psychology from the University of Texas in Tyler. In August 2001 he entered the Graduate School of Biomedical Sciences at the University of Texas Southwestern Medical School at Dallas (UTSWMC) to work on a Ph.D. in clinical psychology. He will graduate from the UTSWMC in August, 2005. He will then begin a two year Fellowship in Clinical Neuropsychology at the Shepherd Center Rehabilitation Hospital in Atlanta, Georgia.

Permanent Address: 8041 Jolly Harbour Court
Wellington, FL 33414