On Diagnosing Alzheimer's Disease: Assessing Abstract Thinking and Reasoning

by

Tamara Elaine Goranson
B.A., University of Saskatchewan, 1991
M.A., University of Saskatchewan, 1993

A Dissertation Submitted in Partial Fulfillment of the
Requirements for the Degree of

DOCTOR OF PHILOSOPHY

in the Department of Psychology

We accept this dissertation as conforming
to the required standard

Dr. Roger Graves, Supervisor (Department of Psychology)

Dr. David Hultsch, Departmental Member (Department of Psychology)

Dr. Helena Kadlec, Departmental Member (Department of Psychology)

Dr. Holly Tuokko, Departmental Member (Department of Psychology)

Dr. Valerie Kuehne, Outside Member (School of Child and Youth Care)

Dr. Grant Iverson, External Examiner (Department of Psychiatry, U.B.C.)

Supervisor: Dr. Roger E. Graves

ABSTRACT

A series of abstract thinking and reasoning tasks was administered to patients with Alzheimer's disease (AD) and a sample of nondemented older adults matched on age, education, and gender variables. The performance of the AD patients was inferior to control subjects on all verbal and nonverbal reasoning tests, including a newly developed test of analogical reasoning, the Goranson Analogy Test (GAT). Preliminary psychometric analyses of the GAT revealed very high internal consistency, good convergent and divergent validity, and adequate predictive validity. Further analyses revealed that reasoning with pictures was just as easy as reasoning with words for AD patients, indicating that modality of presentation has little effect on reasoning performance. Error analyses revealed no qualitative differences in performance between AD patients and nondemented controls. Taken together, the findings suggest that abstract thinking and reasoning abilities decline with the onset of Alzheimer's dementia. A neurocognitive model of analogical reasoning is proposed to account for the study findings.

Examiners:

Dr. Roger Graves, Supervisor (Department of Psychology)

Dr. David Hultsch, Departmental Member (Department of Psychology)

Dr. Helena Kadlec, Departmental Member (Department of Psychology)

Dr. Holly Tuokko, Departmental Member (Department of Psychology)

Dr. Valerie Kuehne, Outside Member (School of Child and Youth Care)

Dr. Grant Iverson, External Examiner (Department of Psychiatry, U.B.C.)

Table of Contents

List of Tables

List of Figures

Acknowledgments

This research was supported by a Training Award from the Alzheimer's Society of B.C.. I am very thankful for their generous financial support.

I gratefully acknowledge the help of my supervisor, Dr. Roger Graves who challenged me to consider current statistical and methodological papers relevant to this project. His insightful questions were always thought provoking, and his comments encouraged me to think critically.

I acknowledge the contributions made by my committee members especially in the early planning stages of this study. I am especially appreciative for the enthusiasm they expressed regarding my newly developed test of analogical reasoning, the Goranson Analogy Test.

I am indebted to the following individuals for their support in helping me to recruit Alzheimer's patients to participate in this study: Sandi Somers, R.P.N. and the Upper Island Geriatric Outreach Program at St. Joseph's Hospital in Comox, B.C.; Drs. Marilyn Bater and Ted Rosenblood, Geriatricians at Homer II Pavilion, Royal Jubilee Hospital; Jenny English, Manager of Victoria Home Care and Hospital Liaison; Dr. B. Lynn Beattie, Geriatrician at the Alzheimer Clinic, Vancouver Hospital & Health Sciences Centre; and, Patty Pitts, Information Officer, University of Victoria.

To all of the Alzheimer's patients who participated in this study. Observing some of their challenges as they attempted to reason analogically was both revealing and puzzling. I am grateful for their courage, wisdom, and willingness to share their time with me.

To Dr. Tuokko who initially sparked my interest in dementia research, aging, and mental health when I completed a practicum under her supervision at Elderly Outreach Services.

To my grandmother, Mary Bodnar, who before her death participated in the pilot testing of the GAT.

To my husband, Doug Chichak, who provided tireless assistance with computer programming, printing, scanning, and other technological assistance. He so unselfishly provided his time and support, encouraging me always to persevere when preparation of that final draft seemed like an unreachable goal.

I appreciate the many gestures of support from family and friends, especially Tiffany, Tom, and Laura who expressed so much interest in my progress and study findings. Thank you for the intellectually stimulating dinner discussions. It also made an important difference to have my brother, John, near me during my internship year in London, Ontario. I appreciate the way he encouraged me to be logical and rational, and to embrace intellectual challenges. His own grammatical expertise and writing are inspirational.

Dedication

I want to dedicate this dissertation to my parents, Alan and Elaine Goranson. Throughout their careers as educators, they modeled a love for learning, teaching me to embrace knowledge and strive for understanding. My dad taught me to face challenges with courage and faith. My mom inspired me to choose instruction and to work conscientiously to find wisdom in my studies and in my life experiences. Thank you both for your encouragement.

On Diagnosing Alzheimer's Disease: Assessing Abstract Thinking and Reasoning

Analogical reasoning is a type of abstract thinking that involves establishing a correspondence between one set of relations and another (Goswami, 1991). Analogical reasoning performance often is studied by presenting problems of the form, "A is to B as C is to D" (or A:B::C:D) with the D term omitted (i.e., **glove** is to **hand** as **shoe** is to **?**). The task is to solve the analogy by choosing the correct D term from among a list of possible answer choices. Traditionally, the analogy problems have been presented in either verbal or nonverbal format.

In verbal analogical reasoning tasks, the items are words or pictures of concrete items whereas in nonverbal tasks, the items typically are geometrical patterns with missing parts. To solve verbal analogies, the participant must choose the correct word or picture that completes the analogy. To solve nonverbal analogies, the participant must choose the correct geometrical puzzle piece that completes the pattern. Both verbal and nonverbal analogical reasoning tasks require the participant to think abstractly about the nature of the relationships governing the terms of the analogy. Thus, analogical reasoning can be considered a type of abstract thinking.

According to the current criteria for the clinical diagnosis of Alzheimer's Disease, abstract thinking and reasoning performance is an important area to assess during neuropsychological testing. NINCDS-ADRDA (National Institute of Neurological and Communicative Disorders and Stroke – Alzheimer's Disease and Related Disorders Association; McKhann et al., 1984) criteria for the diagnosis of *possible* and *probable* AD require that deficits be observed in two or more areas of cognition, one of which must be memory. Although NINCDS-ADRDA criteria do not formally indicate which

other areas of cognition possibly could show deterioration due to the onset of senile dementia of the Alzheimer's type, the criteria require that dementia be established by clinical examination and confirmed by neuropsychological tests (McKhann et al., 1984). Zec (1993) recommends that several areas of cognitive functioning be assessed by the neuropsychologist including language, praxis, visuospatial functioning, attention, problem solving, judgment, and abstract thinking.

Several test batteries for dementia currently include neuropsychological tests of abstract thinking and reasoning [i.e., Canadian Study of Health and Aging (CSHA -2 and 3), 2001; Fuld, 1983; Rosen, 1983; Salmon, & Butters, 1992]. For example, some batteries (i.e., CSHA-3, 2001; Salmon and Butters' University of California, San Diego Alzheimer's Disease Research Center battery, 1992) incorporate the Similarities subtest from the WAIS-R. In this test of verbal concept formation, the participant is required to explain what each of a pair of words has in common. The word pairs range in difficulty from simple ("orange-banana") to more difficult ("fly-tree"). Items are passed at the two-point level if an abstract generalization is given and at the one-point level if a response is concrete. In general, the test requires relational thinking and as such, can be considered a measure of analogical reasoning.

Another commonly employed diagnostic criteria set for identifying Alzheimer's Disease is that established in the Diagnostic and Statistical Manual of Mental Disorders, fourth edition (DSM-IV; American Psychiatric Association, 1994). According to the DSM-IV, the diagnostic criteria for dementia of the Alzheimer's type include the development of multiple cognitive deficits manifested by both memory impairments and deficits in one or more other areas of cognitive functioning. The DSM-IV specifies

which other areas of cognition may be affected in Alzheimer's Disease (AD). Specifically, AD patients might be expected to have deficits in one or more of the following areas: a) aphasia (language disturbance), b) apraxia (motor impairments), c) agnosia (object recognition deficits), and/or d) executive functioning. Executive functioning is a rather elusive term that refers to a number of loosely connected cognitive functions. The DSM-IV identifies planning, organization, sequencing, and abstract thinking as examples of executive functioning abilities.

Although impairments in abstract thinking are listed as one of the possible symptoms of Alzheimer's Disease (AD) in the DSM-IV (American Psychiatric Association, 1994), the ability of patients with AD to think abstractly has been investigated in only a few studies. Of the studies that have been conducted, there are some recent neuropsychological data to suggest that the ability of AD patients to reason abstractly in the early stages of the disease is similar to that of cognitively intact older individuals (Grady et al., 1988). On the other hand, clinicians working in the area of dementia diagnosis strongly suspect that problem-solving skills, abstract reasoning, and decision making abilities are impaired in the early stages of AD (Albert, 1988; Zec, 1993). It is clear that more research is needed to help clarify the discrepancy between the limited empirical data suggesting that reasoning abilities do not decline with the onset of AD and the belief among clinical neuropsychologists that abstract thinking does deteriorate.

The purpose of this study will be to examine both the verbal and nonverbal analogical reasoning performance of cognitively intact older adults and Alzheimer's patients in the early stages of the disease. Such a study would help further our understanding regarding the abstract thinking and reasoning performance of Alzheimer's

Disease patients. In particular, it would be useful for neuropsychological assessment purposes to determine whether or not abstract thinking and reasoning in addition to memory functioning is impaired early in the course of AD. A second purpose will be to determine whether or not there are quantitative and/or qualitative differences in the verbal/nonverbal analogical reasoning performance of Alzheimer's participants in comparison to nondemented older adults. The results of the AD patients' abstract thinking and reasoning performance will be interpreted with reference to a cognitive theoretical model of semantic memory as well as a newly developed neurocognitive model of analogical reasoning.

Abstract Thinking and Reasoning

Verbal Reasoning in Normal Aging

A variety of neuropsychological tests have been devised to assess verbal abstract thinking and reasoning performance. For example, verbal abstract thinking ability has been measured by: 1) the Proverb Interpretation Test (Gorham, 1956) in which the examinee is asked to explain the meanings of proverbs varying in degrees of familiarity; 2) the Similarities subtest of theWAIS-R and the Wechsler Adult Intelligence Scale - Third Edition (WAIS-III; Wechsler, 1997) in which the examinee is required to identify how two words are alike (e.g., apple and banana); 3) the Abstract Words Test (Tow, 1955) in which the examinee is required to discern how two words differ from one another; and, 4) the Visual-Verbal Test (Feldman, & Drasgow, 1951) in which the examinee is shown 24 cards with four objects on each, asked to determine one way in which three of the objects are alike, and then asked to decide how three of the objects are alike in another way.

To date, limited research has been conducted examining the verbal reasoning performance of older adults. However, of the studies that have been conducted, many have shown that verbal abstract reasoning abilities decline with age. Albert (1988), for example, cites research showing that of the subtests comprising the verbal scale of the WAIS-R, performance on the Similarities subtest shows the greatest decline with age. Other researchers have identified age-related decline on the Similarities subtest (Axelrod, & Henry, 1992; Jarvik, 1988; Kaufman, Kaufman-Packer, McLean, Reynolds, 1991).

Not only does performance on the WAIS Similarities subtest decline with advancing age, older adults also show decreased performance on Gorham's Proverbs Test and the Visual-Verbal Test. In one study, these tests were administered to men aged 30 years to 80 years (Albert, Duffy, & Naeser, 1987). Compared to younger men, older men gave significantly more concrete responses when interpreting proverbs; and, they had difficulty thinking flexibly on the Visual-Verbal Test which resulted in lowered overall scores.

While the finding of a significant difference between the performance of the older men and the younger men on these reasoning tasks may suggest age-related decline in abstract thinking and reasoning, other explanations are plausible. According to Salthouse (1992), for example, older adults may perform more poorly than younger adults because they were not exposed to the same educational opportunities as children, and/or because they experienced less cultural stimulation in their environments as their reasoning skills were developing. Based on this explanation, it would be expected that successive generations would perform at progressively higher levels on various verbal reasoning tasks because changes in the external environment contributed to more favorable

conditions for the development of reasoning abilities. According to Salthouse (1992), favorable environmental conditions might positively influence both older and younger adults such that both groups may continue to improve over time without necessarily observing an alteration in the relationship between age and reasoning performance at any given period. Thus, age differences in reasoning performance do not necessarily imply that reasoning abilities deteriorate with advancing age.

Even if there is age-associated decline in reasoning performance, explanations for why increased age is associated with lower levels of reasoning are lacking. According to some (e.g., Cohen, 1981; Salthouse, 2000), the performance of cognitively intact older individuals on reasoning tasks may be worse than that of younger adults because of factors other than declining reasoning capabilities. Older adults, for example, may be able to think abstractly and construct logical inferences, but forget the inferences they constructed due to working memory failures or decreased speed of information processing. According to Cohen, for example, the capacity for abstract thinking and reasoning may remain intact with advancing age but may be negatively impacted by working memory failures associated with remembering or applying relations.

To summarize the literature regarding changes in verbal abstract reasoning abilities in "normal" aging, there is some evidence to suggest that performance on verbal abstract reasoning tasks varies as a function of age, with older adults performing more poorly than younger adults. However, the reasons for these age differences are still poorly understood.

Verbal Reasoning in Alzheimer's Disease. Not only is research related to the reasoning performance of cognitively intact older adults limited, few studies have

investigated the verbal abstract reasoning performance of Alzheimer's patients. To date, there is mixed support for the view that verbal reasoning performance is negatively affected in Alzheimer's dementia. Furthermore, disease severity and dementia type have not been factors that have been systematically controlled. As a result, it has not been established conclusively that a decline in abstract thinking is indeed one of the first hallmarks of the onset of Alzheimer's Disease.

In one of the few studies of the effects of dementia on abstract thinking and reasoning, Heinik and Aharon-Peretz (1993) examined the performance of eleven Alzheimer's patients, ten multi-infarct dementia patients, and nine cognitively intact older people on four abstract thinking measures. The four tasks included: a) proverb interpretation (e.g., Birds of a feather flock together); b) identifying similarities between words and concepts (e.g., table-chair; apple-pear; love-hate); c) identifying differences between words and concepts (e.g., car-train; child-dwarf; airplane-bird); and, d) solving absurdities (e.g., Is my brother-in-law a man or a woman?). Dementia severity levels were not reported in this study.

To analyze the results, Heinik and Aharon-Peretz (1993) had raters judge the abstractness of each participant's answers on a two point scale with `2' being abstract, `1' being partly abstract, and `0' being concrete. Inter-rater reliability co-efficients were calculated and found to be adequate. The mean ratings obtained by the Alzheimer's group, the multi-infarct group, and the cognitively intact control group then were compared.

Although the actual statistics used to compare the ratings were not reported, Heinik and Aharon-Peretz concluded that AD patients obtained lower mean ratings and

therefore had more difficulty than normal elderly controls explaining proverbs, designating the differences between related concepts, and solving absurdities. However, on the Similarities subtest, AD patients and cognitively intact controls obtained similar ratings. Furthermore, the ratings of the AD patients and multi-infarct dementia patients did not differ on any of the four abstract thinking measures administered.

Because these researchers did not report which statistical tests they used to examine the differences between mean ratings, limited conclusions can be drawn regarding effect sizes or even the statistical significance of the observed differences between means. However, these preliminary findings do raise the possibility that only some types of abstract thinking and reasoning may be affected by the onset of Alzheimer's dementia. Furthermore, the finding that patients with different types of dementia performed similarly on all of the abstract thinking measures may suggest that abstract thinking deficits are symptoms of more than one type of dementia. Diagnostically, this may suggest that patients with Alzheimer's dementia can not be distinguished on the basis of the presence of abstract thinking and reasoning deficits. Rather, other dementias would have to be ruled out first.

The results of other studies indicate that performance on the Similarities subtest of the WAIS-R does decline with the onset of dementia (i.e., Hart, Kwentus, Taylor, & Hammer, 1988). In fact, differences as large as three age-corrected scale score points have been reported between a group of mildly demented patients and normal elderly controls (Larrabee, Largen, & Levin, 1985).

Tuokko (1993) also showed that performance on WAIS-R Similarities varies as a function of dementia severity and that those with AD perform worse than cognitively

intact controls on this verbal abstract reasoning task. In this study, the performance of a group of AD patients with either mild, moderate, or severe dementia was compared to the performance of a group of cognitively intact controls on the WAIS-R Similarities subtest. Overall, the mean age-corrected scaled score obtained by patients with mild AD was 3.2 standard deviations below that of normal elderly controls. This finding suggests that even in the early stages of Alzheimer's Disease, verbal abstract thinking and reasoning abilities may be significantly impaired.

On other measures of verbal abstract reasoning, however, AD patients have had only modest amounts of difficulty relative to cognitively intact controls. For example, some studies have examined the ability of AD patients to recognize and interpret proverbs (Andree, Hittmair, & Benke, 1992; Laflech, & Albert, 1995). In general, AD patients were able to understand proverbs, but they often had difficulty explaining them. Explanatory attempts were incomplete and long-winded; however, literal and concrete interpretations were rare. Deficits in proverbial interpretations were attributed to metalinguistic difficulties rather than poor problem solving and abstract thinking abilities.

In another study, the performance of a group of AD patients was compared to the performance of a group of cognitively intact controls on two verbal abstract thinking measures and one visual abstract thinking measure (Lafleche et al., 1995). Only AD patients in the early stages of the disease were included in this study. Patients who obtained a score of 22 or greater on the Mini Mental State Examination, a cognitive screening measure for dementia (MMSE; Folstein, Folstein, & McHugh, 1975), were considered to be in the early stages of the disease process.

The results of this study revealed that reasoning performance, as measured by the Proverb Interpretation Test and the Similarities subtest of the WAIS-R, did not differ between groups. That is, AD patients in the early stages of the disease did just as well on verbal abstract reasoning measures compared to cognitively intact controls. One limitation of Laflech and Albert's (1995) study, however, is that raw score data were analyzed instead of age-corrected scaled scores.

Age-scaled scores permit an examinee's score to be interpreted in relation to the performance of a large standardization group of randomly sampled same-age peers. In the absence of comparing age-corrected scaled scores, the amount of measurement error may be high. That is, there is a greater possibility that test performance is due to factors other than a true deficit. For example, in Laflech and Albert's study, age could have interacted with dementia status to obscure any true deficits attributable to dementia. In particular, if AD patients were younger than the control group, decline in reasoning performance due to age factors for the control group *and* decline due to dementia status for the AD group may have resulted in the observation of no group differences.

While many of the studies reviewed suggest that dementia patients generally have difficulty on verbal abstract thinking and reasoning measures, according to Lezak (1995), some dementia patients do surprisingly well on certain tests such as the WAIS-R Similarities subtest. However, from her perspective, these patients usually possess excellent pre-morbid verbal abilities and well-formed verbal associations, and so their performance does not appear "impaired" relative to less verbally skilled older adults.

In summary, the results of the studies reviewed are somewhat equivocal with respect to whether verbal abstract reasoning is impaired in early AD. While some studies

indicate that performance on some verbal abstract thinking measures does deteriorate with the onset of Alzheimer's dementia, other studies suggest that verbal abstract reasoning abilities do not decline. However, premorbid abilities and age factors were not always taken into account in studies where no deficits were reported. Furthermore, the types of verbal abstract thinking measures administered were not always consistent across various studies.

Some of the studies reviewed revealed that performance deficits among AD patients are sometimes observed on one commonly employed measure of verbal abstract thinking, the Similarities subtest of the WAIS-R. Since performance on Similarities involves a type of reasoning in which the participant is asked to determine how two different words are related, this test provides a very good measure of verbal analogical reasoning. Essentially, the test requires that one be able to think abstractly about the relationships between concepts.

It may be that in early AD, verbal analogical reasoning performance may not be preserved. However, it is possible that some AD patients do poorly on this test because of factors other than reasoning deficits. For example, since the Similarities subtest requires the examinee to retrieve conceptual information from memory (e.g., `apple' and `banana' are both `fruit'), it may be that poor performance reflects impaired memory recall for word associations rather than deficits in reasoning abilities per se.

Nonverbal Reasoning in Normal Aging. One nonverbal analogical reasoning task that is commonly used by neuropsychologists to assess abstract thinking in older individuals is the Raven's Coloured Progressive Matrices Test (RCPM; Raven, Court, & Raven, 1976). The RCPM consists of 36 items divided into three 12-item sets (i.e., Set

A, Ab, and B). The test contains both gestalt completion tasks and some simple visual analogies with the task involving looking at geometric shapes with a missing piece like a puzzle and finding the missing piece from among several answer choices. According to Villardita (1985), items from the RCPM measure visuoperceptual/visuospatial abilities, gestalt-like processing, and analogical and abstract thinking. Lezak (1995) conceptualizes the task as measuring the ability to reason about spatial, design, and numerical relationships as well as offering a measure of fluid intelligence.

To date, few studies have been designed to specifically investigate age-related differences on the RCPM in cognitively intact older adults. However, there is some research (i.e., Anderson, Hartley, Bye, Harber, & White, 1986; Babcock, & Laguna, 1996; Babcock, 1994; Clayton, & Overton, 1976; Cowart, & McCallum, 1984; Koss et al., 1991; Salthouse, & Skovronek, 1992; Schultz, Kaye, & Hoyer, 1980) to suggest that older adults perform less well than younger adults on a more difficult version of the RCPM, the Raven's Advanced Progressive Matrices test (Raven, no date).

Items on the Raven's Advanced Progressive Matrices task are presented in a 3 x 3 matrix with the last cell left blank. To solve the problem, the participant must decide which of eight alternative geometric figures belongs in the last cell, following rules for the rows and columns. Although performance on this task has been shown to vary as a function of age, Babcock and Laguna (1996) specified the nature of the age differences by examining the influence of processing speed, working memory and three hypothesized components including rule induction, rule application, and rule coordination. Examples of rules to be applied to the nonverbal abstract geometrical problems include the addition or subtraction of a smaller figure, or a 90^0 or 180^0 rotation of the original figure.

Babcock and Laguna (1996) found that several cognitive abilities are required to solve the problems presented on the Raven's Advanced Progressive Matrices test. These cognitive abilities include: abstract thinking and reasoning, the ability to manipulate geometric figures according to a given rule, and intact visuospatial/visuoperceptual abilities. According to Babcock et al., performance differences observed between older and younger adults were not due to nonverbal reasoning impairments but rather to difficulties manipulating spatial stimuli.

More recently, Salthouse (2000) examined 11 data sets containing the scores of a large number of cognitively intact adults on various tests of reasoning including the Advanced Progressive Matrices test. Data from samples of between 195 and 390 adults ranging in age from 18 to 80 years of age were examined. His cross-sectional study revealed age differences on the Raven's Advanced Progressive Matrices Test. He also found that in addition to easy items, difficult items were influenced by age.

Only a few studies have examined the effects of age on the Raven's Coloured Progressive Matrices (RCPM). Most of these studies have been designed for the primary purpose of collecting normative data. For example, Orme (1966) collected RCPM normative data for older adults from a standardization sample consisting of 271 people age 60 to 89 recruited from a local medical clinic in Britain, none of whom were suffering from dementia or other mental or neurological disorders. These norms are widely used despite the fact that percentile equivalents are not given for each raw score and age so that the clinician must interpolate intermediate scores and extrapolate more extreme scores (Tuokko, & Hadjistavropoulos, 1998).

In another study (i.e., Measso et al., 1993) the RCPM was administered to 197 cognitively intact people over the age of 60 from six Italian cities. A multiple linear regression was used to examine the proportion of raw score variance attributable to age, gender, and education. On the basis of this regression model, RCPM raw score adjustment values were determined in order to remove the effects of gender, age, and education from the raw score. Unfortunately, Measso and colleagues only report correction values for adults aged 20 to 79, and so these norms are not applicable for use with very old adults.

In addition to these normative studies, there is some research indicating that the greater the age of a person, the poorer their RCPM performance (Heidrich, & Denney, 1994; Panek, & Stoner, 1980). Age effects begin to appear around age 40 (Yeudall et al., 1986) and continue to show a linear decline with advancing age.

Nonverbal Reasoning in Alzheimer's Dementia. Not only is there a paucity of research in the area of nonverbal abstract reasoning and normal aging, little is known about how well dementia patients perform on measures of nonverbal abstract thinking and reasoning such as the RCPM. In one study, Shuttleworth and Huber (1989) presented a subset of problems from the RCPM to three groups of patients; namely, patients with dementia of the Alzheimer's type, patients with pseudodementia, and patients with vascular dementia. The performance of the three patient groups were compared and the results revealed that there were no differences in performance between the number of RCPM problems correctly solved by the AD patients and the other two patient groups. On average, AD patients solved 2.5 out of 6 items correctly. However, no control group was included in this study.

Many of the studies conducted to date have failed to examine the nonverbal reasoning abilities of a matched control group of older adults. As a result, it is difficult to determine the degree to which AD patients have difficulty with nonverbal analogical reasoning problems relative to older adults who are aging "normally". Another limitation of many of the studies conducted to date is that dementia severity has not been systematically taken into account. In Shuttleworth and Huber's study, for example, dementia severity was not reported and could have been a possible confounding variable that differed between groups.

Disease severity has been taken into account in more recent studies where significant deficits among AD patients on nonverbal abstract reasoning measures have been observed only later in the course of the disease (i.e., Christensen, Multhaups, Nordstom, & Voss, 1991; Grady et al., 1988). In a longitudinal positron emission tomography (PET) study of very mildly impaired AD patients, for example, Grady and his colleagues found that one of the first neurological deficits to appear in Alzheimer's Disease is impaired memory for newly learned information. Memory deficits were followed by problems with nonverbal abstract reasoning as measured by the RCPM, problems with attentional processing, impaired visuoconstructive skills, and impaired language functioning. Grady and colleagues found that the timing between each of these sequential neuropsychological impairments is quite variable ranging in length from 1.5 to 6 years, which implies that nonverbal reasoning abilities are not immediately affected with the onset of Alzheimer's disease.

In another much older study (Orme, 1957 in Raven, Court, & Raven, 1990), the performance of a cognitively intact group of older people on the RCPM was compared to

the performance of a group of depressed older adults and a group of dementia patients. All participants ranged in age from 61 to 80 years. Members of the cognitively intact control group in this study were volunteers at a Seniors' Social Club whereas the elderly depressive group and dementia patients were inpatients at a local British hospital. Dementia patients obtained a mean RCPM raw score significantly below that of the other two groups.

Unfortunately, the basis for the diagnosis of dementia was not stated in this study, dementia type was not reported, and dementia severity was not taken into account. As a result, it is difficult to determine whether the performance of AD patients early in the course of the disease would be expected to be worse than the performance of a group of cognitively intact controls on the RCPM.

Diesfeldt (1990) did take into account dementia severity when investigating the RCPM performance of AD patients whose diagnosis was established by clinical examination using NINCDS-ADRDA criteria (McKhann et al., 1984). In this study, the performance of patients with Alzheimer's dementia of moderate severity was compared to the performance of a sample of cognitively intact older adults. AD patients were considered to have moderate levels of dementia if they lacked the capacity for independent living and could not travel to day-care centers on their own. While the purpose of this study was not specifically designed to investigate the nonverbal abstract reasoning abilities of AD patients, adequate statistical methods were used to examine the performance of the two groups on the RCPM.

Overall, Diesfeldt (1990) found that the RCPM performance of AD patients in the middle to late stages of the disease was significantly lower than that of the cognitively

intact controls. However, the question remains: "How do AD patients in the early stages of the disease perform on the RCPM relative to cognitively intact older adults?". While these results certainly suggest that reasoning abilities deteriorate relative to a sample of cognitively intact controls later in the course of the disease, the results can not be applied to AD patients who have recently been diagnosed with dementia.

In another study, the performance of fifteen AD patients and fifteen cognitively intact controls was compared on the RCPM (Raven et al., 1976) as part of a larger study investigating the effects of various encoding conditions on recognition memory performance (Goldblum et al., 1998). Although the degree of dementia severity was not reported, patients who scored under 15/30 on the Mini-Mental Status Examination (i.e., MMSE; Folstein, Folstein, & McHugh, 1975) were excluded from the study and the mean MMSE score was 20.7/30. Traditionally, the cut-off score for dementia on the Mini-Mental State Examination has been 23 or lower (Tombaugh, McDowell, Kristjansson, & Hubley, 1996). In this study, therefore, a mean MMSE score of 20.7 probably indicates that the AD participants included in this study were "mildly to moderately impaired". In terms of the study findings, Goldblum and colleagues found that their AD study participants obtained significantly lower RCPM scores in comparison to the cognitively intact controls.

In summary, only a few studies have been conducted examining the RCPM performance of AD patients; and, most of these studies have been fraught with methodological flaws. As a result, it has not been clearly established that relative to a group of cognitively intact older people, "mildly" impaired AD patients show nonverbal reasoning deficits. While some studies indicate that moderately to severely impaired AD

patients have nonverbal reasoning difficulties, other research indicates that reasoning abilities remain intact early in the course of the disease, but eventually decline as the disease progresses (Grady et al., 1988).

Some research also has been conducted examining whether there are any qualitative differences in nonverbal abstract reasoning performance between dementia groups and cognitively intact individuals on the RCPM. In one study (Raven, Court, & Raven, 1990), for example, RCPM error analyses were conducted for a group of depressed, demented, and cognitively intact older people. Dementia type was not specified, making it difficult to draw conclusions regarding the performance of AD patients, in particular.

Raven and colleagues found that the type of RCPM error most commonly made did not differ as a function of group. The most frequent type of error made by all study participants included choosing an identical appearance match to one of the figural terms in the problem. The next most common error choice involved choosing a figure that was wrongly oriented. While there were no qualitative differences in performance between the depressed, demented, and cognitively intact participants in this study, further studies might be worthwhile to more clearly delineate whether the types of errors made by AD patients, in particular, are unique.

In the event that AD patients can be distinguished from other groups on the basis of their nonverbal reasoning errors, delineating the reasons for these differences would be important. As well, previous research has not yet determined what cognitive processes mediate successful nonverbal reasoning performance. Some researchers (i.e., Diesfeldt, 1990; Salthouse, & Skovronek, 1992) have suggested that in addition to nonverbal

reasoning abilities, visual attention and working memory are needed to complete RCPM analogies successfully.

According to Villardita (1985), other cognitive abilities also are required for successful performance. Villardita (1985) reclassifed the RCPM items and found that eleven RCPM items are contingent on visuo-perceptual ability requiring the participant to identify a match to the figures in the problem that is similar in shape, pattern, and design. Nineteen RCPM items require an analysis of the features characterizing the stimulus (i.e., length and direction of lines, arrangement of dots, wideness of angles, etc.); and, six problems require the discovery of the analogical relationships between the geometrical parts constituting the problem. Thus, according to Villardita, good RCPM performance depends on intact visuo-perceptual ability as well as intact abstract thinking and reasoning abilities.

In summary, it may be that performance on the RCPM requires visual attention, visual-spatial skills, the ability to visually analyze and synthesize information, working memory, and the ability to reason abstractly. Thus, poor performance on the RCPM may not be due to reasoning deficits per se. In the following section, the role that other cognitive components play in both verbal and visual analogical reasoning will be examined in more detail.

Cognitive Skills Required to Reason Analogically

Cognitive theorists have attempted to identify the various information processing subcomponents that are required to reason by analogy in the classical A:B::C:D analogical reasoning paradigm. According to Sternberg's (1977) Componential Theory of Analogical Reasoning, for example, six different cognitive processes are required for

successful verbal analogical reasoning performance. These include: 1) encoding the terms of the analogy and then retrieving from memory a list of attributes associated with each of the terms in the analogy; 2) inferring the association between the A and B terms; 3) discovering the relationship between the A and C terms; 4) finding a relationship between the C and D terms analogous to the relationship shared by the A and B terms; 5) evaluating the goodness of fit of the chosen D term; and, 6) making a response.

The assumption underlying Sternberg's analogical reasoning model is that analogical reasoning performance difficulties can be attributed to difficulties in executing one of the component processes (i.e., encoding terms, accessing and retrieving from memory information related to the terms of the analogy, inference making, or mapping). According to Sternberg (1977), for example, good reasoners more thoroughly encode all of the terms of the analogy compared to poor reasoners.

Sternberg's Componential Theory of Analogical Reasoning was originally devised for adults. However, the model was never used to account for deficits in analogical reasoning performance with advancing age or with the onset of dementia. In fact, experimental support for the model was derived mostly by studying the analogical reasoning performance of children. Goldman, Pellegrino, Parseghian and Sallis (1982), for example, gave 8- and 10-year old children verbal analogies to solve and found that irrespective of age, less skilled responders chose answer options that were highly associated with the C term. They concluded that children who perform poorly on verbal analogical reasoning tasks generally do not encode the A and B terms in the analogy but instead exclusively focus on the C term and choose a D term which is associatively related to the C term.

Gitomer, Curtis, Glaser and Lensky (1987) evaluated how children process verbal analogies by recording both the number of eye fixations and the time that children spent looking at the words comprising the terms of the analogy. Results revealed that the proportion of time spent on the initial processing of each of the individual word terms was greater for the high-ability solvers than the low-ability solvers. In accordance with Sternberg's theory (1977), these results suggest that a more exhaustive initial encoding of the terms of the analogy lead to better analogical reasoning performance.

Linguistic abilities also have been shown to influence verbal analogical reasoning performance (Masterson, Evans, & Aloia, 1993; Nippold, Erskine, & Freed, 1988). In particular, children with expressive and receptive language deficits are less accurate in solving verbal analogies than children without language deficits. Knowledge of the vocabulary items included in the task is obviously an important determinant of good analogical reasoning performance.

Likewise, item familiarity has been found to affect analogical reasoning performance. Goswami and Brown (1989), for example, examined which types of analogical relationships were easiest for children to understand. They found that certain abstractions such as the relation "opposite", the biological relation "habitat", and the causal relation "powered by" were not understood by children as young as 3 years of age. When these same children were asked to reason about items and relationships that they could understand in domains with which they were familiar (e.g., observing pictures depicting the relationship, "one piece of playdough: two pieces of playdough::one piece of apple: two pieces of apple") then they were capable of reasoning analogically.

Several other developmental studies have shown that when very young children are able to access the relevant knowledge base and retrieve from memory conceptual information that is related to the analogy word stems, they demonstrate more successful analogical reasoning performance (Brown, 1989; Brown, & Kane, 1988; Goswami, 1991). Thus, item familiarity, intact object recognition, and knowledge about the relational terms associated with the words or picture stimuli are important to be able to reason effectively.

Although much has been learned concerning the cognitive skills required to reason by analogy in normal and language impaired children, the cognitive skills required to reason by analogy in cognitively intact older adults and adults with dementia have yet to be delineated. However, much can be learned about the cognitive abilities that are required to reason analogically by studying the analogical reasoning performance of young children.

To summarize the developmental literature, children do well on analogical reasoning problems if: a) they encode all analogy terms b) they possess adequate language comprehension skills; c) they are familiar with the terms of the analogy; and, d) they are able to engage in efficient memory retrieval of conceptual information related to the terms of the analogy. The fact that memory retrieval skills are required to perform well on verbal analogical reasoning tasks has direct implications for AD patients who experience semantic memory deficits early in the course of the disease.

Other more recent cognitive models of analogical reasoning (i.e., the Structure Mapping Engine of Falkenhainer et al., 1989; the Analogical Constraint Mapping Engine of Holyoak, & Thagard, 1989; and the Incremental Analogy Machine of Keane, &

Brayshaw, 1988 in Kean, Ledgeway, & Duff, 1994) have expanded on Sternberg's (1977) initial theory of analogical reasoning detailing more specifically the process involved in the mapping stage. Mapping involves infering and applying conceptual relationships. In the classic "A is to B as C is to D" paradigm, it involves determining the correspondence between the source analog (i.e., A and B terms) and the target analog (i.e., C and D terms) (Spellman, & Holyoak, 1996).

According to the Structural Mapping Engine (SME; Falkenhainer et al., 1989) theory, people solve analogies by retrieving all possible conceptual relationships and making all possible matches between the source and target analogs. Among these alternative mappings, the best is selected by using constraints such as favouring alternatives that can map in multiple ways rather than in one way only.

Holyoak and Thagard's (1989) Analogical Constraint Mapping Engine (ACME), on the other hand, conceptualizes analogical reasoning as a process whereby terms access a (memory) network of associational units or nodes. Each node represents a match between two terms. There are excitatory links between nodes, but to enforce a one-to-one mapping, inhibitory connections also exist. According to the ACME model, when the network has been constructed, it is run until the activation settles into a stable state. The nodes whose activation exceeds a threshold corresponds to the optimal set of matches between two domains.

In contrast to Holyoak and Thagard's parallel distributed processing ACME model, Keane and Brayshaw (1988; in Kean, Ledgeway, & Duff, 1994) proposed a serial processing model called the Incremental Analogy Machine (IAM). IAM assumes that a small subset of the possible mappings is generated. If the first mapping that is built is

less than optimal, IAM will undo the matchings found and try to find an alternative mapping that is more optimal and conforms to the constraints of the problem more ideally.

Taken together, the SME (Falkenhainer et al., 1989), ACME (Holyoak, & Thagard, 1989), and IAM (Keane, & Brayshaw, 1988 in Kean et al., 1994) models all attempt to outline how people engage in "mapping", one stage of the analogical problem solving process. The models differ in terms of the number of mappings that are assumed to be generated (i.e., all versus a small subset), and the process by which an ideal mapping is found. One theory stipulates that each mapping is serially tested in order to discover the best relationship between terms. Another model assumes that an ideal mapping is generated based on whether or not it conforms to certain "constraints". The third model assumes that mappings are chosen if they are above threshold in what is described as a network of conceptual associations. Thus, all of the models essentially expand on Sternberg's Componential Theory of Analogical Reasoning by detailing more specifically what happens in step 4 (i.e., finding a relationship between the C and D terms analogous to the relationship shared by the A and B terms).

All three models, however, are limited in the sense that they focus on the mapping stage and fail to account for the importance of other pre-mapping processes essential for successful verbal analogical reasoning performance. For example, according to Gentner (1989), there are several pre-mapping constraints that may limit analogical problem solving, including: 1) an inability to encode the source analog in memory; 2) an inability to retrieve information related to the source analog; and, 3) difficulties selecting subsets

of relational information as inputs to the mapping engine. This latter constraint may involve selective attention failures and/or difficulties inhibiting competing information.

With the exception of Holyoak and Thagard's ACME theory which permits various pre-mapping constraints such as difficulties encoding the source analog in long-term memory, the cognitive models that have been proposed to date have paid little attention to the role that intact memory functioning plays in analogical reasoning. In order to solve a verbal analogical reasoning problem, for example, one must first be able to activate the appropriate conceptual nodes and relevant conceptual relationships in memory. Thus, in addition to being able to perform analogical mappings, access and retrieval from semantic memory is crucial for successful verbal analogical reasoning performance.

Semantic Memory and Alzheimer's Disease

Semantic memory is a memory system that contains knowledge of words, concepts, and their meanings and associations, and it is thought to be disrupted in AD (Butters, Granholm, Salmon, Grant, & Wolfe, 1987; Chan, Butters, Salmon, & McGuire, 1993; Daum, Riesch, Sartori, & Birbaumer, 1996; Grossman, Mickanin, Robinson, & D'Esposito, 1996; Nebes, 1989; Weingartner, Kawas, Rawlings, & Shapiro, 1993). The presence of a semantic memory deficit would seem to place a person at risk for verbal analogical reasoning performance difficulties because solving such problems requires retrieval of conceptual knowledge, knowledge of word meanings, and knowledge of relationships among concepts from memory (Sternberg, 1977).

There is some controversy in the cognitive neuropsychological literature about the extent to which semantic memory is impaired in Alzheimer's disease (Hartman, 1991).

On the one hand, semantic priming studies indicate that the prior presentation of a semantically related word facilitates lexical decision making for both AD patients and cognitively intact older people (Nebes, Martin, & Horn, 1984; Nebes, Brady, & Huff, 1989; Ober, Shenaut, Jagust, & Stillman, 1991; Ober, Shenaut, & Reed, 1995). Intact priming in AD has been taken to indicate preserved semantic memory organization.

On the other hand, other research supports the view that semantic memory disintegrates with the onset of Alzheimer's dementia. For example, AD patients are able to answer questions about an object's superordinate category (e.g., a cat is an animal) but progressively lose knowledge about the object's physical features and functions (Chertkow, Bub, & Seidenberg, 1989; Huff, Corkin, & Growden, 1986; Martin, & Fedio, 1983; Schwartz, Marin, & Saffran, 1979; Warrington, 1975). Since attribute and functional information is thought to be stored in semantic memory, these findings have been taken to indicate that there is deterioration of the semantic memory system itself as a result of the onset of Alzheimer's Disease.

Furthermore, word association studies in which AD patients are given a stimulus word and asked to say the first word that comes to mind have shown a decrease in responses that are semantically related to the stimulus word and an increase in unrelated and perseverative responses (Gewirth, Shindler, & Hier, 1984). These findings also may suggest that semantic memory deteriorates with the onset of Alzheimer's disease.

Other research, however, has shown that AD patients have difficulty accessing and retrieving information from an intact semantic memory store. In one study, for example, AD patients had difficulty describing the use of a presented object but they could pick out from an array of items those items that would be useful in a given chore

such as cooking dinner (Flicker, Ferris, Crook, & Bartus, 1987). In other studies, AD patients have demonstrated an ability to correctly sort a series of pictures of objects into appropriate categories supplied by the experimenter even though they could not generate exemplars of the categories spontaneously (Martin, & Fedio, 1983; Huff et al., 1986; Weingartner, Kawas, Rawlings, & Shapiro, 1993). Furthermore, when given a concept in the form of a word or picture, AD patients could identify the category that the concept belonged to (e.g., A cat is an animal). However, AD patients could not spontaneously generate category examples (e.g., "Tell me the names of some animals") (Ober, Dronkers, Koss, Delis, & Friedland, 1986).

More recent findings are consistent with the view that AD patients have a relatively intact semantic memory (Bonilla, & Johnson, 1995). When AD patients in the early stages of the disease were asked to sort cards representing one of two categories (e.g., animals or occupations) in a way that reflected their similarity to each other, AD patients were able to use their semantic knowledge to make appropriate groupings. In fact, the AD patients' sorting schemes matched that of controls although the AD patients were less able to clearly explain their rationale for the groupings.

Taken together, the results of these studies indicate that the organization of semantic memory may remain relatively intact with the onset of Alzheimer's disease. In some cases, however, Alzheimer's patients experience difficulties accessing and retrieving information from this intact semantic memory store. The inability to access and/or retrieve information from semantic memory would imply that verbal analogical reasoning, which requires the ability to spontaneously retrieve concepts and associations from semantic memory, also would be impaired in Alzheimer's Disease. Less clear is

whether or not nonverbal analogical reasoning would be impaired in Alzheimer's Disease.

The Structure of Semantic Memory

In cognitive neuropsychology there is an extensive debate about whether the structure of semantic memory is amodal or modality specific (Shallice, 1993). According to some, there is only one semantic memory system that is accessed equally by picture input and words (Potter, & Faulconer, 1975; Caramazza, Hillis, & Rapp, 1990). Another view is that there are two separate semantic memory systems - one for pictures called "visual semantic memory"; and, one for words called "verbal semantic memory" (Pavio, 1991; Riddoch, Humphreys, Coltheart, & Funnell, 1988). Visual semantic memory is assumed to store functional and perceptual information associated with picture stimuli and verbal semantic memory is assumed to store functional and perceptual information associated with word stimuli.

Whether or not semantic memory is amodal or modality specific has direct implications for reasoning tasks presented in either word or picture format. If, for example, semantic memory is modality specific and if AD patients have difficulty accessing one of the two semantic memory stores or if one of the two memory stores is degraded, then reasoning performance may vary as a function of modality of presentation.

Is semantic memory modality specific? One early study found support for the hypothesis that semantic memory is modality specific. Warrington's (1975) patient EM had a diffuse dementing illness that resulted in impaired semantic memory. When EM was asked questions about the perceptual attributes and functional uses of objects, he

made more errors when the objects were presented in word format than when the objects were presented in picture format. Warrington accounted for these performance differences by proposing that EM had a degraded verbal semantic memory but intact visual semantic memory.

Similar dissociations have been observed by other researchers. In one study, for example, patient MP showed very poor comprehension of written and spoken words but had no difficulty answering questions regarding the attributes and superordinate category membership of pictures (Bub, Black, Hampson, & Kertesz, 1988). Another patient who had primary progressive aphasia was unable to understand words referring to animals, but experienced no difficulty understanding and describing the conceptual attributes of these same animals displayed in pictorial form (McCarthy, & Warrington, 1988). For example, when the patient was presented with a picture of a "rhinoceros", he could identify that the animal lived in Africa. However, when the word "rhinoceros" was presented, the patient could not identify the animal's natural habitat.

Traditionally, researchers have accounted for these dissociations by hypothesizing that neurological insults selectively impair one of the two semantic memory stores. In contast, other researchers postulate that there is only one semantic memory system where perceptual attribute information as well as functional information is stored for both word and picture stimuli. To explain the modality effects, these researchers postulate that all components of semantic memory remain intact, but only certain components are accessible.

In a single case study, for example, Chertkow, Bub and Kaplan (1992) asked an AD patient to name pictures of animals and then to match pictures of these same animals'

heads to pictures of their corresponding hind quarters. Only pictures of animals were presented not the corresponding word labels. Thus, if visual semantic memory was impaired due to the onset of Alzheimer's disease, neither perceptual attribute information nor the names of the pictured animals should have been accessable. Surprisingly, Chertkow and colleagues found that the AD patient was unable to name any of the pictures of the animals in full. However, the patient was able to carry out the front-to-back matching task correctly. It was assumed that the matching task could not be performed without accessing the part of semantic memory where perceptual attribute information was stored. Thus, Chertkow and colleagues argued that AD patients were only able to access the part of semantic memory that holds stored details regarding an object's physical attributes, but that other information stored in semantic memory was inaccessable.

The results of this latter experiment raise the possibility that modality effects could be observed in Alzheimer's disease because of a disconnection between various component parts of semantic memory rather than a deterioration of the component parts, themselves. Further support for this theory comes from another experiment comparing Alzheimer's and control participants' ability to answer three types of questions about objects presented in either word or picture format (Chertkow, Bub and Kaplan; 1992). Specifically, AD patients and controls were asked superordinate category membership questions (e.g., "Is this a tool or clothing?"), perceptual attribute questions (e.g., [item=saw] "Is the edge made of metal or wood?"), and functional questions (e.g., [item=saw] "Is it used on a piece of wood or on a stone?"). Superordinate category questions were answered equally well by AD patients and controls. AD patients, on the

other hand, made more errors than controls answering perceptual attribute questions and functional questions for both word and picture stimuli. Chertkow and colleagues (1992) concluded that impaired access to certain components of semantic memory occurred for both word *and* picture stimuli.

Farah and McClelland's (1991, 1992) Parallel Distributed Processing (PDP) Model of Semantic Memory offers a theory regarding the nature of the component parts constituting semantic memory. The model also can be used to account for any observed modality effects influencing the reasoning performance of Alzheimer's patients. This model will be reviewed in the next section.

Parallel Distributed Processing Model of Semantic Memory

According to Farah and McClelland's PDP model of semantic memory, the basic architecture of semantic memory is a parallel distributed processing network where pieces of information are represented at nodes in a highly interconnected associative network. There are three pools of units in the model: 1) verbal name units (i.e., word labels to enter semantic memory); 2) picture units (i.e., pictorial inputs to enter semantic memory); and, 3) one central semantic memory unit.

One of the assumptions of the Parallel Distributed Processing (PDP) model is that separate verbal and visual semantic memory stores do not exist. Rather, there is one semantic memory unit containing "visual units" and "functional units". The visual units consist of perceptual attribute information and descriptive information regarding the appearance of objects and living things. Functional units, on the other hand, contain nonvisual descriptions of either metaphorical verbal associations (e.g., [item=lion] "king of the jungle") or functions of the item (e.g., [item=cow] "animal that supplies milk").

An additional assumption of the model is that there are bidirectional connections between units, with the exception that there are no direct connections between the name and picture units. A schematic diagram of the PDP model of semantic memory is shown in Figure 1.

To test some of the assumptions of their model, Farah and McClelland used their PDP model of semantic memory to simulate the behavior of McCarthy and Warrington's (1988) patient who could identify pictures of living and nonliving things, names of nonliving things, but not names of living things. Farah and McClelland hypothesized that if the name units were disconnected from the visual semantic units in their PDP model, the model could account for the differences in this patient's ability to identify living and nonliving things presented in different modalities.

Farah and McClelland assumed that irrespective of whether nonliving things are presented as name units or pictorial units, nonliving things primarily are known by their functional features and therefore activate functional semantic memory units before visual semantic memory units. A second assumption was that living things primarily are known by their visual features and therefore activate visual semantic memory units before functional semantic memory units irrespective of mode of presentation.

Given these assumptions, disconnecting the name unit from the visual semantic memory unit was expected to impact the identification of living things presented in word format only. This is exactly what happened with McCarthy and Warrington's patient who was unable to identify the names of living things. The patient could identify pictures of living things because of intact connections between picture units and visual semantic memory. The names of nonliving things were recognized via the route from name units

to functional semantic memory whereas pictures of nonliving things were recognized via the route from picture units to functional semantic memory. Thus, the PDP model was able to account for this patient's performance by focusing on disconnections between different units in the system.

The PDP model also can account for partial disconnections between units in semantic memory. The idea is that if degraded connections between units exist, inappropriate patterns of activation are created in semantic memory which interfere with incoming input arriving from other semantic memory entry routes. For example, if there was a partial disconnection between the name units and visual semantic memory, the primary route of entry into semantic memory for a word depicting a living thing would be greatly, though not completely, blocked. As a result, semantic memory would be accessed via the connection between the name unit and the functional semantic memory unit.

According to the assumptions of Farah and McClelland's PDP model, the few degraded connections between the name unit and the visual semantic memory unit that do exist would produce inappropriate patterns of activation in the visual semantic unit. As a result, collateral connections from functional semantic memory would be unable to activate the correct patterns in the visual semantic memory component and errors in object identification would be made.

Despite the complexity of the PDP model of semantic memory (Farah, & McClelland, 1991, 1992), the model is useful in that it underlines the importance of considering that modality effects may be due to disconnections rather than underlying semantic memory impairments. That is, neurologically impaired patients with semantic

memory access or retrieval deficits may demonstrate word and pictorial dissociations due to a myriad of disconnections or partial disconnections between semantic memory units.

Memory and Verbal Analogical Reasoning

It has been hypothesized that effective verbal analogical reasoning performance depends on successful retrieval of conceptual information from memory. In other words, the integrity of functional semantic memory and the ability to access and retrieve information from this memory system are necessary for successfully analogical reasoning performance. It is assumed that access to the component of semantic memory where visual attribute information is stored is less important for successful verbal analogical reasoning performance because in order to solve verbal analogies, inferring abstract relationships between concepts is more important than comparing perceptual and physical attribute information.

If one adopts Farah and McClelland's PDP model of semantic memory, it could be hypothesized that performance on verbal analogies critically depends on the integrity of the connections to functional semantic memory. For word problems, this would mean that the connections between the name units and functional semantic memory would have to be intact. For picture problems, the connections between the pictorial units and functional semantic memory would have to remain intact.

With the onset of Alzheimer's Disease, poor performance on verbal analogical reasoning problems in which the terms of the analogy were words could be explained by a partial but not a full disconnection between the name units and functional semantic memory units. A full disconnection would not explain performance deficits on word problems because words could still be encoded into semantic memory via the connection

between the name units and visual semantic memory and then via the connection between visual semantic memory and functional semantic memory. A partial disconnection between the name units and functional semantic memory, on the other hand, could explain any performance deficits on verbal analogies in word format.

According to the assumptions of Farah and McClelland PDP model, the degraded connections between the name units and functional semantic memory would produce inappropriate patterns of activation in the functional semantic units. These inappropriate patterns of activation would in turn interfere with the ability of collateral connections from the visual semantic memory units to activate the correct patterns of activation in functional semantic memory.

In contrast, poor performance on verbal analogies presented in pictorial format could be explained if the connections between the pictorial units and functional semantic memory were partially degraded. In this case, degraded connections between the pictorial units and functional semantic memory would produce inappropriate patterns of activation in functional semantic memory. These inappropriate patterns of activation would in turn interfere with the ability of collateral connections from the visual semantic memory units to activate the correct patterns of activation in functional semantic memory.

In summary, Farah and McClelland's PDP model of semantic memory can be used to account for any performance differences between reasoning analogically with pictures compared to words in early AD. Several assumptions, however, have to be adopted in order to use the model to explain the verbal analogical reasoning performance of AD patients. First, one would need to assume that access to functional semantic

memory is crucial for successful analogical reasoning performance. Second, the observation of either a pictorial advantage or a word advantage must be assumed to result from partial disconnections between units rather than complete disconnections. Third, a pictorial advantage is assumed to result from a partial disconnection between name units and functional semantic memory. Fourth, a word advantage is assumed to result from a partial disconnection between pictorial units and functional semantic memory.

Justification for a Pictorial Advantage in Analogical Reasoning

Although the verbal and nonverbal analogical reasoning performance of AD patients has not formally been studied, the performance of AD patients on other types of verbal and nonverbal reasoning tasks has been examined. In one recent study, for example, patients with mild to moderate Alzheimer's disease as well as a control group of older adults were asked to complete several measures of cognition including verbal and visual-spatial reasoning tests (Christensen, Multhaup, Nordstrom, & Voss, 1991). For the verbal reasoning task, all participants were required to choose from four names of objects the one that was most similar to two other objects. For the visual-spatial problem solving task, all participants were required to choose which of four figures would be made if a set of geometric pieces were put together.

For both the verbal reasoning task and the visual-spatial problem solving tasks, the AD patients performed significantly worse than the control group of cognitively intact older adults. Furthermore, AD patients performed equally poorly on both the verbal reasoning task and the visuospatial problem solving task.

Although there is some limited research (i.e., Christensen et al., 1991) to suggest that modality of presentation does not affect the reasoning performance of AD patients,

the results of other studies (e.g., Bub, Black, Hampson, & Kertesz, 1988; McCarthy, & Warrington, 1988; Warrington, 1975) suggest that modality of presentation does exert an influence on other tasks. These studies have generally shown that AD patients are better able to retrieve information from semantic memory related to pictorial input. Given this research, AD patients may be expected to do better on an analogical reasoning task in which the terms of the analogy are pictures instead of words.

Justification for a pictorial advantage can be modeled in Farah and McClelland's PDP semantic memory model. In particular, if there is a partial disconnection between the name units and functional semantic memory early in the course of the disease, AD patients would be expected to perform better when presented with picture analogies rather than word analogies. Pictures would simultaneously access both visual and functional semantic memory via intact connections. Words, on the other hand, would be unable to fully and directly activate functional semantic memory except via the connections between the name units and visual semantic memory.

If degraded connections between the name units and functional semantic memory caused an inappropriate pattern of activation in functional semantic memory, there would be interference with collateral connections arriving from visual semantic memory. As a result, the appropriate relationships needed to solve the verbal analogy in functional semantic memory would not be activated and performance on verbal analogies presented in word format would be poor.

For cognitively intact older adults who are aging "normally", on the other hand, no semantic memory disconnections or partial disconnections are expected. As a result, cognitively intact seniors should perform equally well when the terms of the analogy are

either words or pictures. Word terms would be expected to enter semantic memory most directly via the connection between the name units and functional semantic memory. Pictorial terms would be expected to enter semantic memory most directly via the connection between the pictorial units and functional semantic memory. No interference from any other partially disconnected routes of entry would be expected in someone who is aging "normally".

Study Goals and Hypotheses

The primary goal of this study was to examine the abstract thinking and reasoning performance of Alzheimer's patients relative to a matched sample of nondemented controls to determine whether or not verbal and nonverbal reasoning abilities remain intact early in the course of the disease. In this study, abstract thinking and reasoning performance was measured using both traditional reasoning measures as well as a newly developed test of analogical reasoning. The new test was called the Goranson Analogy Test (GAT) and it included both word and picture analogies. The psychometric properties of the GAT were examined in this study to determine whether the test offers a reliable and valid measure of verbal analogical reasoning for use with older adults.

The purpose for developing the GAT was threefold. First, the inclusion of both picture and word analogies on the GAT will allow for an assessment of any modality effects influencing the reasoning performance of AD patients. Second, the development of a standardized neuropsychological measure of verbal abstract reasoning may offer yet another test of executive functioning to include in a dementia assessment battery. Third, although some reasoning tests assessing various aspects of verbal reasoning (i.e., logical thinking, sequential reasoning, concept formation, practical judgments) have been

developed, there are few neuropsychological tests of analogical reasoning available. Of those that have been developed [i.e., Matrix Analogies subtest from Kaufman Assessment Battery for Children (Kaufman, & Kaufman, 1983); Luria's Concept Formation Test, 1966 in Lezak, 1995; Opposite Analogies subtest from the Stanford Binet Intelligence Scale (Terman, & Merrill, 1973)], none have had enough neuropsychological use with older adults to result in published studies.

In terms of the study predictions, based on the limited research data available, there is some evidence to suggest that dementia related nonverbal abstract thinking and reasoning deficits should not be expected in the early stages of Alzheimer's Disease. Nonverbal analogies are assumed to offer a more "pure" measure of abstract thinking and reasoning compared to verbal analogies in which access to conceptual knowledge in functional semantic memory is required. Thus, if reasoning abilities per se are intact early in the course of Alzheimer's disease, then Alzheimer's patients should perform just as well as a matched sample of cognitively intact controls on all nonverbal reasoning measures.

On the other hand, verbal abstract thinking and reasoning requires the ability to retrieve conceptual knowledge from semantic memory. Given the fact that semantic memory access and retrieval deficits are associated with the onset of AD, Alzheimer's patients should perform more poorly compared to the nondemented control group on all verbal reasoning tests, including the GAT. Without the ability to access retrieve conceptual relationships, verbal abstract thinking and reasoning performance should be impaired.

Based on a review of the literature, there is also some limited evidence to suggest that AD patients can access semantic memory more easily when pictorial stimuli compared to word stimuli are presented (Bub et al., 1988; McCarthy et al., 1988; Warrington, 1975). Thus, Alzheimer's patients should perform better when presented with GAT picture analogies compared to word analogies. A pictorial advantage can be explained using Farah and McClelland's PDP model of semantic memory if all of the connections between the pictorial units and functional semantic memory are still intact but the connections between the name units and functional semantic memory are degraded. For nondemented older adults, no differences between reasoning with pictures and reasoning with words are expected if access and retrieval from the functional component of semantic memory is preserved in normal aging.

A final goal of this study is to examine whether or not there are any qualitative differences in the analogical reasoning performance of Alzheimer's patients compared to a matched sample of nondemented controls. To examine this question, the types of errors that AD patients and nondemented controls make on both verbal and nonverbal reasoning tasks will be examined. For the nonverbal reasoning tasks, errors will be examined to determine whether or not AD patients choose answer choices by matching the visuoperceptual attributes of the answer choice to either the B term or the C term of the analogy. As well, the error type most frequently made by the AD group will be compared to the error type most frequently made by the control group to determine whether or not there are any group differences in performance. Such information may provide qualitative evidence for differences in the reasoning strategies that AD patients and nondemented controls use when solving nonverbal analogies.

For the GAT - a verbal analogical reasoning test - error analyses also will be conducted. For both the picture and the word analogies, five different answer choice options will be presented including the correct answer, a strong associate to the C term, a category match to the C term, a mere appearance match to the C term, and an unrelated answer choice. If AD patients have difficulty inhibiting the tendency to respond associatively to the C term of the analogy instead of processing the whole problem and reasoning analogically, then the answer choice most frequently chosen to complete the problem should be the erroneous strong associate. Alternatively, if AD patients focus on the surface features or perceptual attributes of the C term of the analogy, then the answer option most frequently chosen should be an appearance match error.

A qualitative analysis of the errors made by AD patients and nondemented older adults also will be made to determine whether or not the type of error most frequently made differs between groups. This type of information would be useful to help clinicians and researchers understand any information processing differences that exist between AD patients and nondemented older adults when solving verbal analogies. For example, if the error most commonly made by the AD patient is an appearance match error but the error most commonly made by the control group is a category match error, then different hypothesis regarding semantic memory activation in normal aging versus Alzheimer's dementia can be made.

Method

Participants

Sixty nondemented older individuals and forty-nine Alzheimer's (AD) patients residing in the province of British Columbia, Canada participated in this study. The control

group of nondemented older adults was recruited through local newspaper advertisements, advertisements at seniors' centers and local church groups, and referrals from other study participants. Each nondemented participant was paid $12.00 to participate in the study. AD patients each were paid $20.00 for study participation.

Geriatricians, neurologists, neuropsychologists, and geriatric nurse specialists practicing in hospital settings in Victoria, Comox, and Vancouver, British Columbia, Canada referred the greatest percentage of AD patients (i.e., 67%). Four percent of the AD sample were recruited from care facilities in Greater Victoria, six percent were referred by professionals affiliated with the provision of Home Support Services in the Greater Victoria area, and twenty-three percent responded to advertisements. Of those responding to the advertisements, 3 patients had already taken part in other university studies for which they had been required to submit confirmatory diagnostic information. Three other patients responded to advertisements placed at the Alzheimer's Resource Centre of the Capital Regional Health District of Victoria, British Columbia, Canada; and, 5 patients responded to local Victoria newspaper advertisements. For those who responded to the newspaper advertisements, the name of the physician who had made the diagnosis was obtained and consent from the patient to contact this physician was sought to confirm the diagnosis.

Overall, 69% of the Alzheimer's sample resided in Victoria, B.C. and 23% resided in the Comox Valley Region. The remainder of the AD sample (i.e., 8%) either lived in the Vancouver area, the Gulf Islands Region, or smaller communities around Duncan, B.C.. All of the control group participants were residents of the Greater Victoria region of Vancouver Island.

Prior to study participation, all of the AD patients had received a formal diagnosis of *possible* or *probable* Alzheimer's Disease on the basis of either NINCDS-ADRDA (McKhann et al., 1984) diagnostic criteria or DSM-IV diagnostic criteria (American Psychiatric Association, 1994) depending on the referral source. Diagnoses were made based on the evaluations of a number of different professionals including neurologists, psychiatrists, neuropsychologists and/or geriatricians. All AD patients had been administered appropriate laboratory tests (i.e., computed tomographic scan, blood chemistry, total thyroxine, folate, and B_{12} levels serology, urinalyses) to rule out alternative causes of dementia such as tumour, stroke, or metabolic causes.

All participants in both the AD group and the control group were English-speaking and were self-assessed to be able to read and write which was later verified by observing performances on a cognitive screening measure, the *Mini-Mental State Examination* (MMSE; Folstein, Folstein, & McHugh, 1975). The ability to hear and see also was grossly assessed by observing each individual's performance during mental status testing with close attention paid to any difficulties listening to and following test instructions, and/or difficulties seeing sufficiently well enough to copy a geometrical design and/or read a sentence. On this basis, all patients were assessed to have vision and hearing sufficient for compliance with testing procedures.

Participants were excluded from study participation if their history suggested evidence of previous neurological dysfunction (i.e., past history of brain tumour, head injury, stroke, or epilepsy), psychiatric illness, exclusionary medical conditions (i.e., diabetes, cardiovascular disease, oncologic disorders of recent onset), and/or alcohol and drug addiction problems. Participants who were being treated for hypertension and for

hypothyroidism were included for data analyses only if their conditions were reported to be well controlled with medications and only if these co-morbid medical conditions had not caused any neurological complications.

The control group of participants (n = 60) was screened for dementia using the *Modified Mini-Mental State Examination* (3MS; Teng, & Chui, 1987) which includes the *Mini-Mental State Exam* (MMSE). Both the 3MS and MMSE are widely used cognitive screening measures comprised of a number of subtests measuring orientation to time and place, rudimentary language skills, praxis, memory, and attention and concentration. The 3MS differs from the MMSE in that it includes four additional subtests. Specifically, participants are asked to: (a) provide the date and place of their birth; (b) generate as many animals they can think of in 30 seconds; (c) identify the similarities between word pairs; and, (d) recall three previously learned words after a time delay. The 3MS is scored out of 100 and the MMSE is scored out of 30.

Traditionally, scores below 78/100 on the 3MS and scores below 23/30 on the MMSE are considered to be indicative of cognitive impairment (Tombaugh, & McIntyre, 1992). However, it recently has been suggested that the use of cut-scores to gauge the presence or absence of dementia is problematic because the sensitivity and specificity of these cognitive screening measures vary as a function of age and education (Tombaugh, McDowell, Kristjansson, & Hubley, 1996). As a result, Tombaugh et al. (1996) recommended that raw scores be converted to standardized scores using their normative data, broken down by age and education levels.

In this study, MMSE and 3MS raw scores were converted to percentile scores based on the norms provided by Tombaugh et al. (1996). According to Lezak (1995), the lower

limit of the percentile range corresponding to "low average" falls at the 9th percentile. "Borderline impaired" scores range from the 3rd to the 8th percentile. Any score less than the 3rd percentile is considered to be impaired.

The control group's (n = 60) MMSE percentile scores ranged from the 11th percentile (low average) to the 98th percentile (superior). The mean MMSE percentile score for the control group (n = 60) fell at the 57th percentile (sd = 25.3). On the 3MS, the control group's percentile scores ranged from the 20th percentile (low average) to the 100th percentile (very superior), with the mean falling at the 70th percentile (sd = 22.9). Thus, no one in the control group obtained a score in the borderline or impaired range on these cognitive screening measures for dementia. On this basis, it was concluded that this sample of participants could be considered "nondemented".

In addition to administering the MMSE to the control group of participants, the MMSE also was administered to the AD group. The MMSE is widely used as a rough indicator of dementia severity (Reisberg et al., 1986; Tombaugh, & McIntyre, 1992) with lower scores thought to be indicative of greater cognitive impairment. Only patients who fit the criteria for early or "mild" Alzheimer's Disease were included in this study.

Previous researchers traditionally have classified early dementia using MMSE cut-scores rather than percentile scores. The problem with this method is that cut scores vary as a function of age and education (Tombaugh et al., 1996). That is, the cut score for an old-old person with limited education may be higher than the cut-score for a young-old person with post-secondary education. One limitation of previous research, therefore, is that cut-scores have been rigidly set without systematically examining the demographic characteristics of the sample under study.

Furthermore, there has been wide variability in terms of which cut-score has been set as a marker for "mild" or "early" dementia. For example, some studies (i.e., Hughes, Graham, Patterson, & Hodges, 1997) use MMSE scores as low as 16/30 to classify AD patients as "mildly demented". This corresponds to percentile scores lower than the third percentile for all age and education levels in Tombaugh et al.'s (1996) normative sample. Reisberg and his associates (1986) specified that a MMSE score range from 10 to 19 constitutes the most definitive indication of "early dementia". Other researchers (Nicholas, Obler, Au, & Albert, 1996) have designated MMSE cut-off scores < 21 to be indicative of mild AD. Still others (Galsko et al., 1990; Landerman, Blazer et al., 1991 in Tombaugh, & McIntyre, 1992; Sabe, Jason, Juejati, Leiguarda, & Starkstein, 1993; Welsh, Butters, Hughes, Hohs, & Heyman, 1991) classified patients who obtained MMSE scores between 18 and 23 as "mildly demented".

In the present study, mild AD was determined on the basis of percentile score consideration not cut-scores. Identifying mild AD in this way was judged to offer an opportunity to exclude those cases in the more advanced stages of AD not afforded with the more popular method of using a fixed cut-off criterion. MMSE percentile scores were derived with reference to Tombaugh et al.'s (1996) normative data stratified for age and education. Only those AD patients whose MMSE percentile scores were equal to or greater than the 3rd percentile were included in the final data analyses. AD patients obtaining percentile scores lower than this (i.e., scores falling in the severely impaired range) were assumed to be in the later stages of dementia and so their data were not analyzed. No upper limit percentile score restriction was imposed because the results of

some studies indicate that patients with mild forms of dementia can obtain high MMSE scores (i.e., Galasko et al., 1990; Sabe et al., 1993)

In this study, mild dementia was operationally defined as a clinical diagnosis of possible AD plus a MMSE percentile score greater than or equal to the 3rd percentile. On the basis of these criteria, two AD patients' data were eliminated. Both of these patients' MMSE percentile scores fell at the 1st percentile, indicating that these patients were most likely in the moderate to severe stages of the disease. Thus, the total number of AD participants whose data were analyzed was 47. The AD group's MMSE percentile scores ranged from the 3rd percentile (borderline impaired) to the 66th percentile (average). The mean MMSE percentile score for the AD group (n = 47) fell at the 16th percentile (sd = 14.9).

Both AD patients and nondemented controls were excluded from study participation if they scored in the severely depressed range (i.e., greater than 19) on the *Geriatric Depression Scale* (GDS; Yesavage et al., 1983). The GDS is a brief self-report measure for depression consisting of 30 questions to be answered in "Yes/No" format. The GDS provides a measure of common affective and behavioral symptoms associated with depression with scores of 0-9 being "normal", scores of 10-19 indicating "mild depression", and scores of 20-30 indicating "severe depression". In this study, the mean GDS score for the control group (n = 60) was within the normal range at 4.3 (sd = 3.13; range: 0 - 14). The mean GDS score for the AD group (n = 47) was 7.1 (sd = 3.78; range: 1-14), which also is within normal limits.

The *Boston Naming Test* (BNT; Kaplan, Goodglass, & Weintraub, 1983) also was administered to all participants to screen for object recognition deficits (i.e., agnosia). The

BNT is a 60-item test traditionally used to measure confrontation naming. In this study, the BNT was administered to ensure that participants could recognize simple objects well enough to be presented with analogical reasoning problems in which the terms of the analogy were pictures. Thus, participants were eligible for study participation even if they performed poorly on the BNT because of anomia. If, on the other hand, they performed poorly because of the presence of agnosia, their data were excluded from further analyses.

To assess the presence of agnosia using the BNT, the participant was first asked to name the pictured object. If the participant was unable to name the object, he/she then was asked to identify the use of the object or what was known about the object. An ability to identify functional attribute information was assessed to indicate that object recognition was intact. For all of the participants tested, no one was found to have significant object recognition impairments. That is, participants could always identify the correct use of the object even if they were unable to name the object.

For all study participants, raw BNT scores were converted to percentile scores on the basis of normative data provided by Ross, Lichtenberg and Christensen (1995). These norms are based on naming abilities not object recognition abilities. Ross et al.'s norms were used because: (a) their normative sample had similar demographic characteristics compared to the group of older adults who participated in this study; and, (b) they provide norms which are broken down by age and years of education. The range of BNT percentile scores obtained by both the AD group and the control group is presented in Figure 2.

Figure 2 shows that the BNT percentile scores of the AD patients were highly variable with some patients performing in the impaired range and others demonstrating above average performance (range:1st percentile to 86th percentile). As noted previously,

even when AD patients had a severe confrontation naming impairment (BNT score = < 3rd percentile), their object recognition abilities were still judged to be intact. That is, they usually could spontaneously provide stimulus cue information (i.e., the functional use of the object; the category the object belonged to) without being prompted, or if they did require prompting, they often were able to generate the name of the object. Thus, even though they could not name the object, they were judged not to have any difficulties correctly perceiving the objects.

Figure 2 also shows that the range of percentile scores for the nondemented sample was positively skewed. Thus, in terms of confrontation naming ability, most nondemented participants were performing within or above the average range in comparison to a sample of same-aged peers with similar education levels. Just like the AD sample, those nondemented participants who obtained scores below the average range had confrontation naming impairments. However, their object recognition abilities were judged to be intact.

Matched samples. Of the 60 cognitively intact participants, 47 were matched to the Alzheimer's group on the basis of age, education, and gender. Age and education were matched to within 4 years. These matched group sizes provided a statistical power of .79 to detect an assumed moderate effect size (d =.5) for observed p-values < .05 (Cohen, 1992).

Demographic data for the 47 AD participants and the matched group of 47 nondemented participants are presented in Table 1. As can be seen from Table 1, after matching, there were no group differences on age or education (t (92) = .535, p = .594 and t (92) = - 1.448, p = .151, respectively).

Premorbid estimates of verbal IQ (VIQ) also were examined to ensure that the Alzheimer's patients and the matched group of nondemented participants had similar verbal

abilities. Since some of the tasks administered in this study involved verbal reasoning and verbal concept formation, it was important to match the verbal abilities of the two groups as closely as possible to prevent this factor from being a potential confound. Tests of premorbid Verbal IQ were used because of time considerations related to administering standardized psychometric measures of intelligences. As well, a "premorbid" estimate of verbal IQ was obtained to account for any verbal intellectual decline associated with the onset of Alzheimer's dementia. The intent was to ensure that the verbal ability of the AD group prior to their developing dementia was as high as that of the control group.

Two methods of determining premorbid VIQ were employed. First, all participants were administered the *National Adult Reading Test - Revised* (NART; Blair, & Spreen). For this test, the examinee is required to read out loud a list of irregularly pronounced words ranging in degrees of difficulty. The number of errors made in pronunciation is computed and then verbal IQ is determined by entering this score into a regression formula developed in 1992 by Ryan and Paolo [i.e., Estimated VIQ = 132.3893 + (NART errors) (- 1.164)]. This equation stands up well on cross validation studies with cognitively intact elderly people, meaning that there is good VIQ predictive accuracy with an independent sample and with a 5-year delay between measurement of VIQ and the administration of the NART (Carswell, Graves, Snow, & Tierney, 1997).

A second means of predicting VIQ is by using demographic-based formulas. Use of demographic-based regression equations are advocated because demographic variables, unlike reading ability measures, are resistant to cognitive decline associated with a dementing disease process (Graves, Carswell, & Snow, 1999). On the other hand, early on in the course of a dementing illness, reading tests such as the NART do resist the effects of

mild AD well (Stebbins, Wilson, Gilley, Bernard, & Fox, 1990) and are more accurate predictors of IQ than are demographic predictors (Graves et al., 1999). One commonly used demographic-based formula is the Barona equation (Barona, Reynolds, & Chastain, 1984)[1]

In the current study, premorbid estimates of VIQ were calculated using both the NART and the Barona Index for the AD group and the group of matched nondemented control group participants. Mean VIQ premorbid estimates as determined using the NART and Barona Index are presented in Table 2. Analyses revealed that the mean VIQ premorbid estimate for the AD group (estimated VIQ mean = 105.2, sd = 9.26) did not differ from that of the matched control group (estimated VIQ mean = 107.94, sd = 8.26) on the basis of the Barona Index (t (92) = - 1.497, p = .138). This finding would be expected given that the two groups were matched on some demographic variables that were also used in the Barona equation to estimate premorbid VIQ.

On the other hand, using the NART to estimate premorbid VIQ, the AD group was found to have slightly lower estimated VIQ scores (estimated VIQ mean = 106.3; sd = 12.35) compared to the control group (estimated VIQ mean = 111.34; sd = 11.52) of nondemented participants (t (92) = - 2.05; p = .043). Because the NART is only reasonably resistant to cognitive deterioration, the matching was not questioned and the AD group was

[1] VIQ = 54.23 + .49 (age) + 1.92 (sex) + 4.24 (race) + 5.25 (education) + 1.89 (occupation) + 1.24 (residence)

Note: Age: 55-64 = 7; 65-69 = 8; 70-74 = 9.
Sex: male = 2; female = 1.
Race: Black = 1; other ethnicity = 2; white = 3.
Education: 0-7 years = 1; 8 = 2; 9-11 = 3; 12 = 4; 13-15 = 5; 16+ = 6.
Residence: urban = 2; rural = 1.
Occupations: Professional and technical = 6; managers, officials, proprietors, clerical and sales workers = 5; craftsperson and skilled workers = 4; not in the labour force = 3; operatives, service workers, farmers and semiskilled workers = 2; farm labourers and unskilled workers = 1.

believed to possess relatively similar premorbid verbal intellectual abilities compared to the control group.

Table 3 shows the mean performances of the AD group and the matched nondemented group on the following screening measures: a) the *Geriatric Depression Scale* (*GDS*; Yesavage et al., 1983); b) the *Boston Naming Test* (*BNT;* Kaplan, Goodglass, & Weintraub, 1983); and c) the *Modified Mini Mental State Exam* (*3MS*; Teng, & Chui, 1987) which contains the *Mini Mental State Examination* (*MMSE*; Folstein, Folstein, & McHugh, 1975).

Analyses revealed that the AD group had a higher mean GDS score compared to the matched sample of nondemented controls ($d = .75$; $t\ (92) = 3.74$; $p < .001$). However, a slightly elevated GDS score for the AD patients in comparison to the nondemented control participants was expected given that a number of the symptoms listed on the GDS may occur as normal concomitants of the dementia process. For example, when an insightful AD patient affirms the presence of memory difficulties and problems with attention, these symptoms may reflect the patient's health status rather than his or her emotional status.

Even though the mean GDS scores differed between groups, both means still fell within the "normal" range indicating that neither the AD group nor the nondemented group was reporting symptoms typically associated with a clinically significant depression. Furthermore, the range of scores for the two groups was similar, with the highest score for both groups falling in the "mildly depressed" range. Depression in this range will not impact the cognitive test performance of Alzheimer's patients (Raskin, 1986) nor will it impact the performance of cognitively intact older adults (Niederehe, 1986) or patients being assessed for suspected cognitive impairment (Bieliauskas, 1993). Therefore, the

presence of mild levels of depression in some AD and nondemented participants was not considered a significant factor that would influence performance on abstract thinking and reasoning measures.

Table 3 also shows the mean Boston Naming Test percentile scores for the AD group and the matched sample of nondemented controls. As expected, t-test analyses indicated that the mean BNT percentile score of the AD group was significantly lower than the mean BNT percentile score of the nondemented sample ($d = 1.6$; $t(92) = -6.1$, $p < .001$). However, this difference was attributed to the fact that some of the AD patients were anomic, a factor believed to have little influence on analogical reasoning performance.

Materials

The test battery for both the nondemented participants and the Alzheimer's patients consisted of four screening measures, two verbal abstract thinking measures, and two nonverbal abstract thinking measures. The four screening measures included cognitive status tests [i.e., *Modified Mini Mental State Exam* (*3MS*); *Mini Mental State Examination* (*MMSE*)]; a self-report mood questionnaire [i.e., *Geriatric Depression Scale* (*GDS*)]; a confrontation naming test [i.e., *Boston Naming Test* (*BNT*)]; and, a reading test designed to measure premorbid verbal IQ [i.e., *National Adult Reading Test - Revised* (NART)]. These tests have been described previously.

The two verbal abstract thinking measures used in this study included: (a) a measure traditionally used by neuropsychologists to assess verbal reasoning abilities; and, (b) a newly developed test of verbal analogical reasoning. The conventional measure was the Similarities subtest from the WAIS-III. For this test, participants are presented with word pairs ranging in difficulty from easy (*orange - banana*) to more difficult (*fly-tree)* and are

asked to explain how the words in each pair are similar to one another. Responses are scored in accordance with an answer key in which typical responses are listed. Abstract responses are given 2 points, concrete or partially abstract answers are given 1 point, and incorrect responses are given no points.

The second verbal reasoning measure administered in this study was a newly developed test designed to measure verbal analogical reasoning ability: the Goranson Analogy Test (GAT). Two versions of the GAT were created for the purposes of this study each containing 24 word analogies and 24 picture analogies. GAT (Version 1) is presented in Appendix A. GAT (Version 2) is presented in Appendix B.

For each version of the GAT, the analogy problems were presented in the form: "A is to B as C is to D" (i.e., A:B::C:D), with the D term omitted. For each problem, five answer choice options to complete the D term of the analogy were shown below the problem matrix. Answer choice options consisted of four incorrect answer choices and one correct answer choice. Incorrect answer options were designed so that no matter what answer was chosen, something could be determined about the thinking processes underlying the choice. For example, incorrect answer choices corresponded to the C term of the analogy in the following manner (e.g., [problem: **cheese:cow::egg:?**]: one was a strong associate answer (option = **bacon**), one was a category match (i.e., **milk**), one was a mere appearance match (i.e., **sun**), and one was an unrelated object (i.e., **camel**). Error classification labels for each problem on GAT (Versions 1 and 2) are presented in Appendix C.

Error classifications in this study were similar to the ones used by Goswami (1991) and Goswami and Brown (1989) who studied the errors made by children on analogical

reasoning problems. Even though the analogical reasoning problems used in this study were modeled after those of Goswami and Brown (1990), each test item was original in terms of stimuli content and the uniqueness of the analogical relationships depicted.

The two versions of the GAT differed from one another in terms of modality of presentation. That is, the concepts presented in pictorial format on GAT (Version 1) were presented in word format on GAT (Version 2). Thus, the two versions of the GAT were not parallel forms of the test because the modality of concept presentation differed.

Two nonverbal abstract thinking and reasoning measures also were administered in this study: (a) the Raven's Coloured Progressive Matrices Test (RCPM); and, (b) the Matrix Reasoning subtest from the WAIS - III. On the RCPM, the examinee is presented with a visual pattern with a missing piece like a puzzle and asked to select from six answer choices the item that best completes the visual pattern. The RCPM includes 36 problems and the total score is the number of items solved correctly. The RCPM items also can be classified in terms of three problem groupings (Villardita, 1985). Set I includes 11 items contingent on the use of the principle of sameness or visuo-perceptual ability (i..e., A_1, A_2, A_3, A_4, A_5, A_6, Ab_1, Ab_2, Ab_3, B_1, B_2). Set II includes 19 items contingent on the use of the principle of symmetry (i.e., A_7, A_8, A_9, A_{10}, A_{11}, A_{12}, Ab_4, Ab_5, Ab_6, Ab_7, Ab_8, Ab_9, Ab_{10}, Ab_{11}, B_3, B_4, B_5, B_6, B_7). To solve these problems correctly one must be able to assess the length and direction of lines, the arrangement of dots, and the wideness of angles. Set III includes 6 problems contingent on the discovery of analogical relationships (i.e., Ab_{12}, B_8, B_9, B_{10}, B_{11}, B_{12}).

The WAIS-III Matrix Reasoning subtest includes a series of incomplete patterns that the examinee completes by pointing to or saying the number of the correct response from

five possible answer choices. There are 26 items to complete and there is no time limit. In addition to measuring nonverbal analogical reasoning, correct performance also requires pattern completion and sequential reasoning.

Procedure

With the exception of two AD patients who were tested as hospital inpatients, AD patients and nondemented control group participants were tested individually in their homes after obtaining informed consent from the patient. Each testing session lasted approximately two hours. Most participants were visited only once, although some of the AD patients and/or their families requested that testing be conducted in two one hour sessions.

Each of the neuropsychological tests was administered in accordance with the standardized testing instructions described in the administration manuals. The instructions for the administration of the GAT also were standardized. Briefly, each participant was given two practice problems one of which was in picture format and one of which was in word format (see Appendix A & B). Practice problems were the same across GAT versions, and the modality of presentation did not differ. For the practice problems, coaching was provided on how to solve the problems, correct answers were provided in the event of incorrect responses, and a rationale for why that particular answer choice was correct was given. Before beginning the GAT, participants also were reminded to examine each answer choice before deciding which answer option best completed the analogy. No time limits were imposed. Specific test instructions for the GAT are presented in Appendix D. The approximate time required to complete the 48 item GAT, irrespective of which version was administered, ranged from 20 minutes to 40 minutes, with AD patients being observed to require longer to complete the test compared to nondemented controls.

Half of the AD patients and half of the nondemented controls received GAT (Version 1) and the other half received GAT (Version 2). There were no demographic differences between the group of participants who were administered GAT (Version 1) and the group of participants who were administered GAT (Version 2). Specifically, the two groups did not differ on age (t (92) = .565, p = .573) or education (t (92) = 1.91, p = .06). The performance of GAT (Version 1) examinees did not differ from the performance of GAT (Version 2) examinees on a measures of depression [i.e., GDS (t (92) = 0, p = 1)], on measures of cognitive impairment [i.e., MMSE (t (92) = .299, p = .76); 3MS (t (92) = -453, p = .65)], and on a measure of confrontational naming [i.e., BNT (t (92) = -1.171, p = .25)]. Estimated VIQ levels for GAT (Version 1 and 2) participants also did not differ [i.e., NART (t (92) = .558, p = .58)].

Each participant received 48 unique analogies. Twenty-four of these problems were picture analogies and twenty-four were word analogies that did not correspond to the pictorial concepts. For each version of the GAT, the order in which the two sets of problems (i.e., word versus picture) were administered was counterbalanced across participants. The order in which the problems within each set (word or picture) were administered also was adjusted to control for fatigue effects by beginning with problem #1 for participant #1, and problem #2 for participant #2, etc. Additionally, the sequence in which all of the verbal and nonverbal reasoning measures that were administered in this study was counterbalanced.

Results

Scoring

All of the standardized neuropsychological measures were scored in accordance with the procedures outlined in each of the test administration manuals. For the nonverbal analogical reasoning measures, raw scores on the Matrix Reasoning subtest were converted to age-scaled scores on the basis of WAIS-III (Wechsler, 1997) norms. Raw scores were converted to percentile scores for the Raven's Coloured Progressive Matrices Test (RCPM) on the basis of Orme's (1966) norms. The normative database from which Orme (1966) derived his norms consisted of 271 people age 60 to 89 drawn from a medical clinic visited by people over age 65 residing in Britain.

For the verbal analogical reasoning measures, raw scores on the WAIS-III Similarities subtest were converted to age-scaled scores on the basis of the Wechsler (1997) norms. For the GAT, raw scores were used. For each version of the GAT, participants were awarded 1 point for each correct response for a maximum of 24 points on the picture analogies and 24 points on the word analogies. Mean picture and mean word analogy scores were calculated separately for the AD group and the cognitively intact control group. Unless otherwise mentioned, an alpha level of .05 was used for all statistical tests.

Analyses of the Psychometric Properties of the GAT

Item Unidimensionality. One important assumption of measurement theory is that the various items forming a test all measure one thing in common. That is, all test items should be adequately homogeneous or unidimensional. There is no one

universally accepted method for assessing the homogeneity of a test; however, an evaluation of a test's homogeneity can be made by comparing the percentage of variance explained by the first and second components in a principal component analysis (Cooke, & Michie, 1997). Traditionally, when the ratio between the first and second component is approximately 3:1, the test is considered to have adequate homogeneity. The implication is that the larger the amount of variance explained by the first component, the closer the set of items is to being unidimensional. Another method recommended by Reckase (1979 in Hattie, 1985) is that the first component should account for at least 20% of the variance.

Four separate principal component analyses were conducted on the word and picture problems for each version of the GAT. Data from the combined sample of AD participants ($\underline{n} = 47$) and nondemented controls ($\underline{n} = 60$) were used for these analyses. Separate analyses were conducted to control for possible modality effects affecting factor loadings. That is, since it was not theoretically possible to determine whether or not mode of presentation (i.e., word versus picture analogies) was a variable affecting the unidimensionality of the test items, to be cautious, the items were separated into four clusters. These clusters included: GAT Version 1 word problems; GAT Version 1 picture problems; GAT Version 2 word problems; and GAT Version 2 picture problems.

For each principal component analysis, the performances of the individual items were first scrutinized. Items not loading highly on the first component were eliminated until each principal component analysis provided evidence of unidimensionality. Additionally, an attempt was made to eliminate the same items on GAT (Versions 1 & 2). This left 19 word problems and 19 picture problems for each version of the GAT.

Table 4 lists the GAT items that were eliminated and the eigenvalues and explained variance for the first two factors. In every case, the first factor not only accounted for over 20% of the variance, but it also accounted for 2 to 3 times the variance explained by the second factor. These results indicate that the items that were retained on each version of the GAT were adequately unidimensional. Data from these newly revised 38-item versions were used in the subsequent analyses. The mean, standard deviation, and range of scores obtained by the AD and nondemented groups for the four 19 item sections of the GAT plus the two 38 item GAT versions are presented in Table 5. Figure 3 provides a schematic diagram of the distribution of GAT scores for the AD and nondemented groups.

Once the problematic items had been eliminated, 19 word problems and 19 picture problems on each version of the GAT remained. To examine the unidimensionality of each version of the GAT without separating the problems on the basis of modality of presentation, further principal component analyses were conducted using data from both groups combined. For GAT Version 1 (38 items), the principal component analysis revealed that the first factor had an eigenvalue of 9.43 and accounted for 25% of the total variance, whereas the second factor had an eigenvalue of 2.9 and accounted for 8% of the total variance (ratio of 3:1). These results indicated that the test was essentially unidimensional. For GAT Version 2 (38 items), the principal component analysis also showed adequate unidimensionality. The eigenvalue ratio was 2.8: 1 and the first factor accounted for the greatest amount of the variance (i.e., 23% versus 8%).

In summary, the results of the principal component analyses indicated that the

items forming each version of the GAT were unidimensional. This suggests that all test items were presumably measuring one underlying factor - verbal analogical reasoning ability. As such, an assessment of the unidimensionality of this new analogical reasoning test provided a measure of the GAT's content validity. Item unidimensionality is also an important factor affecting the reliability of a test or the consistency of test scores over time. If all of the items on the GAT assess one underlying trait - verbal reasoning ability - then the GAT score should measure verbal analogical reasoning comparably on separate occasions.

Reliability. An assessment of the reliability of a test is important in order to establish the consistency of test scores over time or the accuracy and stability of test scores across situations (Anastasi, & Urbina, 1997). Traditionally, several methods have been used to assess a test's reliability. These include: inter-rater reliability, test-retest reliability, alternate forms reliability, and methods to assess the internal consistency of the test items. Due to the objective scoring procedures of the GAT, inter-rater reliability of the GAT was not investigated. It was not possible to determine the test-retest reliability of the GAT in this study because the same version of the GAT was not completed by the same study participants on multiple occasions. Alternate forms reliability could have been computed considering that parallel forms of the GAT were developed. However, each participant did not complete both versions of the GAT.

In this study, therefore, the reliability of the GAT was assessed by examining the internal consistency of the word and picture problems on each version of the GAT. Internal consistency reliability is concerned with the interrelatedness of a set of test items. Cronbach's (1951) coefficient alpha is one method of assessing the internal consistency

of a test. Coefficient alpha is measured by calculating the average internal correlation between pairs of items on a test after only one test administration. High coefficient alpha suggests that test items are interrelated but not necessarily unidimensional (Schmitt, 1996).

The traditional cutoff point for acceptable internal consistency as measured by coefficient alpha is .80 (Dijkers, 1997). Using this criteria, GAT (Version 1) showed adequate internal consistency (combined groups: $\alpha = 91$, AD: $\alpha = .81$, nondemented: $\alpha = .77$). The internal consistency of GAT (Version 2) also was acceptable (combined groups: $\alpha = .90$, AD: $\alpha = .84$, nondemented $\alpha = .88$).

Another method for assessing internal consistency is by using the split-half correlation technique. To determine the split-half correlation of the two versions of the GAT, the tests were each divided on the basis of odd even items. The two test halves then were correlated. An estimate of the reliability of the entire test was then computed by using the Spearman-Brown formula. Similar to the results of the coefficient α analyses, the results of this reliability analysis revealed that both versions of the GAT offered reliable measures of analogical reasoning ability (i.e., combined groups: $r = .92$, AD: $r = ..805$, nondemented: $r = .77$ for GAT Version 1; and, combined groups $r = .90$, AD: $r = .835$, nondemented: $r = .88$ for GAT Version 2).

The usual clinical reason for testing is to estimate a person's true score. The *true score* is the score that a person would obtain if there were no measurement errors (Nunnally, 1978 in Charter, 1996). The higher the reliability of a test, the greater the likelihood that the obtained score is close to the true score. Thus, unreliable measures

yield test scores with large measurement errors whereas perfectly reliable measures provide an exact indication of a person's true score.

To estimate the amount of error inherent in the GAT, the standard error of estimate (SE_E) for each GAT version was calculated. The standard error of estimate refers to the distribution of true scores if an obtained score is held constant (Charter, 1996). The SE_E is commonly calculated with the formula: $SE_E = SD_x(r_{xx}(1- r_{xx}^2))^{1/2}$, where *SD* is the standard deviation of the test and *r* is the reliability coefficient of the test. In the case of GAT (Version 1), the SE_E for the combined groups equalled 2.9 (AD = 3.5, nondemented = 2.5) using the reliability coefficients from the coefficient α analyses. For GAT (Version 2) the SE_E for the combined groups was 3.1 (AD = 3.4, nondemented = 2.9).

The standard error of estimate is the appropriate error term to use to determine the confidence interval around an estimated true score (Charter, 1996). The 90% confidence interval is found by the following: estimated true score plus or minus 1.65 times SE_E (where 1.65 is the z score for the .10 level). Confidence intervals for tests with high reliability have narrower bands than confidence intervals for tests with low reliability. Given that both versions of the GAT have relatively high reliabilities and consequently relatively low standard error of estimates, the 90% confidence intervals would be narrower than tests with lower reliability. The interpretation is that we can be sure that approximately 90 people out of 100 people with a particular obtained score will have true scores within an interval still falling within the average range as opposed to intervals spanning across several classifications (i.e., "Impaired" to "High Average").

Validity. To determine the validity of the GAT as a measure of analogical reasoning ability for use with older adults, two questions must be answered. Does the test measure abstract thinking and reasoning? How well does it measure abstract thinking and reasoning? In answering these questions, three types of validity are discussed: content validity, construct validity, and criteria-related validity.

The GAT was judged to have adequate content validity on the basis of the following considerations. First, GAT problem content was presented in analogical reasoning format (i.e., A:B::C:D). According to Lezak (1995), this type of test calls for relational reasoning and abstract comparisons between concepts to be made. Therefore, analogy problems on the GAT were judged to provide a sensitive measure of reasoning ability and abstract thinking. Second, the GAT incorporated a number of different analogical relationships commonly referred to in the analogical reasoning literature (e.g., "comes from"; "used to make"; "needs"; "is used by"; "is made by"; "lives in"; "is made of"; "uses", etc.). Since the item content was not restricted to one type of analogical relationship, the content validity of the measure was judged to be adequate. Third, the analogies on the GAT were presented in two different modalities (i.e., picture versus word analogies). By varying the mode of presentation, the GAT was judged to adequately represent the various kinds of analogical reasoning problems that have been developed to date.

Construct validity, or the degree to which the GAT measures reasoning and abstract thinking, was established by: (a) showing that the GAT was correlated with other tests designed to measure abstract thinking and reasoning (convergent validity); and,

(b) by showing that the GAT was more weakly correlated with a test not purported to measure reasoning ability (divergent validity).

Table 6 shows that both versions of the GAT were significantly correlated with other measures of abstract thinking and reasoning, but uncorrelated with a measure of mood (i.e., Geriatric Depression Scale). As can be seen in Table 6, the correlations were significant at either the .05 or the .01 level. Furthermore, both versions of the GAT were moderately to highly correlated with all of the other reasoning measures. The intercorrelations between the RCPM, Matrix Reasoning, and Similarities subtest also were quite similar to the correlations involving the GAT. This finding also supports the assertion that all of these tests are measuring a similar construct.

On the other hand, the finding that neither version of the GAT was correlated with the GDS offers strong support for the divergent validity of the GAT. Essentially, these results show that the construct that the GAT was designed to measure is quite distinct from that of the GDS. Furthermore, none of the other reasoning measures was correlated with the GDS. Therefore, the GAT performed much the same way as other measures of abstract thinking and reasoning, but unlike a measure of mood.

The criteria-related validity of each version of the GAT also was evaluated. Criteria-related validity is the effectiveness of the test in determining an individual's performance on a criterion measure. One type of criteria validity is predictive validity or the degree to which the test scores are able to predict the criterion under study. In this case, the "criterion" was the diagnosis of Alzheimer's dementia. Thus, the predictive validity of the GAT was the extent to which each version of the test was useful in identifying AD patients and nondemented individuals. In order to determine how

effective the GAT was at discriminating between disease positive and disease negative people, the sensitivity and specificity of each version of the GAT was examined at various cut scores.

Sensitivity refers to the ability of a test to correctly identify those individuals who have been classified as having Alzheimer's dementia according to some generally accepted criteria or gold standard (i.e., DSM-IV; NINCDS-ADRDA criteria, clinical judgment). In this study, therefore, sensitivity was the proportion of AD participants who had GAT scores in the test-positive range (i.e., GAT score fell at or below a given cut score). Specificity was the proportion of disease-negative people (i.e., nondemented participants) with scores falling in the test-negative range (above a given cut-score). That is, specificity refers to the GAT's ability to correctly identify those participants who previously had been classified as cognitively intact.

According to many neuropsychologists, a good clinical test is ideally both highly sensitive and specific at a certain cutoff score. Optimally, high sensitivity is most desirable for screening tools (Essex-Sorlie, 1995 in Tuokko, & Hadjistavropoulos, 1998) in order to ensure that no cases of the disease are missed prior to a more extensive follow-up evaluation. Because one of the purposes of this study was to evaluate whether or not the GAT is a good diagnostic marker of the onset of Alzheimer's Disease, an attempt was made to choose a cut-off score that would balance the sensitivity and specificity of the GAT.

Test specificity allows one to determine the degree to which using the test would result in older people being misclassified as having dementia when they truly are cognitively intact (i.e., false positive errors). Moreover, neuropsychologists using the

results of such a test might spend considerable time conducting an in-depth assessment on people who are not demented if the specificity of the GAT was poor.

Using the 38-item total GAT score, the sensitivity and specificity values were computed at various cut scores for each version of the GAT. The data from the matched samples were used to determine the sensitivity and specificity of each version of the GAT. Table 7 shows the sensitivity and specificity values at select cut-off scores on GAT (Version 1) and GAT (Version 2).

As can be seen from the table, sensitivity and specificity values were highest at a cut score of 27/38 for GAT (Version 1) and at a cut score of 28/38 for GAT (Version 2). On GAT (Version 1), 80% of Alzheimer's patients and 86% of the nondemented sample were correctly classified using GAT scores less than 27 as indicative of impairment. For GAT (Version 2), however, use of this same cut-score resulted in only a moderately high sensitivity and specificity values. For GAT (Version 2), balanced sensitivity and specificity values were obtained at a cut score of 28/38.

Information on the sensitivity and specificity of the other reasoning measures administered in this study also was examined at various cut-scores. Table 8 lists the sensitivity and specificity values associated with the RCPM, the WAIS-III Matrix Reasoning and Similarities subtests, and both versions of the GAT. Overall, the GAT appeared to be equally sensitive compared to the other abstract thinking and reasoning measures. The GATs' specificity, on the other hand, was better than the specificity of the other measures. Thus, both versions of the GAT performed equally well if not better than the other tests of abstract thinking and reasoning in terms of being able to discriminate between groups.

A related and perhaps more critical clinical question is how well does a positive test score (i.e., scores below 27 or 28 on GAT Versions 1 & 2, respectively) predict the presence of Alzheimer's dementia? Table 9 lists the positive predictive values of both versions of the GAT, the RCPM, and the WAIS-III Matrix Reasoning and Similarities subtests. The positive predictive power of a test is defined as the ratio of correctly identified positive cases to the total number of positive cases (i.e., a/a + b). That is, of all those people who obtained a score in the test-positive range, the proportion of people who were correctly diagnosed as having AD is defined as the positive predictive power of the test. Inspection of the positive prediction data in Table 9 reveals that relative to the other tests, the GAT had higher positive predictive values. Thus, a GAT score in the test positive range was more strongly associated with the diagnosis of Alzheimer's dementia, indicating that this newly developed test is a worthwhile test of dementia.

In addition to examining the positive predictive value of the GAT, the likelihood ratio also was examined to help interpret scores. The likelihood ratio is the probability that a given test result would be expected in a patient with AD compared to the probability that the same result would be expected in a nondemented patient (Tombaugh, & McIntyre, 1992). If someone obtains a GAT score of 21, for example, what is the likelihood that that individual truly has AD? Likelihood ratios also are helpful in interpreting a GAT score for an individual from a particular sample, provided the prevalence of dementia is known for the sample (Tombaugh, & McIntyre, 1992). Likelihood ratios therefore allow one to determine the extent to which administering the GAT would improve the pre-test probability of diagnosing Alzheimer's Disease.

The sensitivity and specificity data associated with the GAT were used to derive a likelihood ratio (i.e., sensitivity/(1 - specificity). On GAT (Version 1), the likelihood ratio for a positive test result was 5.7. On GAT (Version 2), the likelihood ratio for a positive test result was 2.56. Thus, if a patient obtained a score below 27/38 on GAT (Version 1), this result would be approximately six times as likely to be seen in someone with AD as opposed to someone without AD. On GAT (Version 2), however, scores below 28 would be approximately three times as likely to be seen in someone with AD compared to someone who was not demented.

In this study, the pre-test probability of Alzheimer's Disease was 50 percent because half of the sample had AD. Did administering the GAT improve the pretest probability of detecting AD? To address this question, the post-test probability of detecting AD was calculated. The calculation formulas and relevant statistics for determining the post-test probabilities of the GAT (Versions 1 and 2) are presented in Table 10. As can be seen from the table, the probability of detecting AD after administering the GAT (Version 1) increased by 36% (i.e., from 50% to 86%). In contrast, by administering GAT (Version 2), the probability of detecting AD increased by 20% (i.e., from 50% to 70%).

Table 11 lists the likelihood ratios, pre-test probability, and post-test probabilities of the RCPM and WAIS III Matrix Reasoning and Similarities subtests. Overall, all of the abstract thinking and reasoning measures administered in this study increased the posttest probability of detecting AD. However, the GAT performed the best at increasing the post-test probability of detecting AD. GAT (Version 1), in particular, was shown to

be a worthwhile diagnostic test in that it increased the probability of detecting AD by a moderately high amount.

Summary. Overall, the psychometric analyses of each version of the GAT revealed that these newly developed verbal analogical reasoning measures have adequate reliability and validity for assessing the abstract thinking and reasoning abilities of a matched sample of AD and nondemented study participants living in the Greater Victoria area. Both versions of the GAT are unidimensional and have adequate high internal consistency and split-half reliability estimates. In addition, both versions were judged to have adequate content validity. Further analyses revealed that the GAT correlated well with other reasoning measures but not with a measure commonly used to assess an older adult's mood status. As such, the GAT showed good convergent and divergent validity.

Finally, the GAT's predictive validity for assessing the presence of Alzheimer's dementia was reasonably high at a cut-score of 27/38 for GAT (Version 1). While GAT (Version 2) was sensitive at detecting AD at a cut score of 28/38, it had low to moderate levels of specificity. Thus, by using a cut-score of 28 on GAT (Version 2), one would run the risk of making a relatively high number of false positive errors (i.e., diagnosing AD in someone who is truly not demented). Further analyses revealed that both GAT versions performed equally well if not better than the other measures of abstract thinking and reasoning in terms of increasing the pre- to post-test likelihood of detecting AD. Both GAT versions also did relatively well at predicting the presence of impairment indicating that these newly developed tests are useful as diagnostic instruments.

Verbal Abstract Thinking and Reasoning Results

GAT analyses. To determine whether or not verbal abstract thinking and reasoning abilities decline with the onset of Alzheimer's disease, the total GAT scores of the AD group ($n = 47$) were compared to those of a matched sample of nondemented participants. It was predicted that AD patients would have verbal abstract reasoning difficulties on account of semantic memory retrieval deficits. Thus, it was expected that the total GAT mean score of the AD group would be lower than that of the nondemented sample irrespective of GAT version type.

A second goal was to determine whether or not modality of presentation affects reasoning performance. Since there is some limited research to suggest that AD patients can retrieve conceptual information from semantic memory more easily when the information is presented in pictorial format, a GAT pictorial advantage was expected for AD patients. For the nondemented group, picture and word analogies were expected to be equally difficult. Modality effects were examined by breaking down the total GAT scores for each participant into a separate score on the word problems (maximum score = 19) and a separate score on the picture problems (maximum score = 19). Mean word and picture scores for each version of the GAT then were computed.

To test all of the study predictions, a 2 x 2 x 2 ANOVA in which there were two test versions (GAT Version 1 & GAT Version 2), two groups (AD participants and nondemented participants), and two analogy types (word analogies and picture analogies) was conducted. Since this involves a total of seven effects (3 main, 4 interactions), a Bonferroni correction was used to control the overall type I error rate. Thus, a $p < .007$

alpha level (.05/7) was used. Figure 4 shows the mean GAT picture and word problem scores as a function of group membership and GAT version type.

The Version x Group x Analogy Type interaction was nonsignificant ($F_{(1, 90)}$= 6.8, MSE = 29.72, p = .01). Neither the Group by Version Type ($F_{(1, 90)}$= 2.7, MSE = 52.37, p = .103) nor the Group x Analogy Type ($F_{(1, 90)}$= .322, MSE = 1.45, p = .332) interactions were significant. Consistent with study predictions, there was a significant Group main effect ($F_{(1, 90)}$= 46.16, MSE = 892.6, p < .001) with AD patients performing worse than the nondemented group. The effect size for this Group effect, as indexed by Cohen's d was d = 1.29, a large effect according to Cohen's criteria (Cohen, 1992). The Analogy Type main effect also was significant (d = .76; $F_{(1, 90)}$= 35.36, MSE = 154.5, p <.001). However, contrary to expectations, a pictorial advantage was not observed.

Further analyses showed that the Version x Analogy Type interaction was significant ($F_{(1, 90)}$= 33.03, MSE = 144.32, p<.001) indicating that the pattern of performance on word and picture analogies changed depending on what version of the GAT was administered. For GAT (Version 1), the Analogy Type effect was not significant ($F_{(1, 45)}$= .024, MSE = 45.0, p = .877). Thus, on GAT (Version 1), word and picture analogies were equally difficult for both AD and control group participants. For GAT (Version 2), the Analogy Type effect was significant ($F_{(1, 45)}$= 58.01, MSE = 45.0, p <.001). However, both groups of participants performed better on the word analogies compared to the pictorial analogies. The effect size for this Analogy Type effect was small according to Cohen's (1992) criteria (i.e., d = .25)

WAIS-III Similarities subtest analyses. For the WAIS-III Similarities subtest, the mean age-scaled scores of the AD patients were compared to the mean age-scaled scores

of the nondemented participants using t-tests. Figure 5 shows the mean age-scaled score that the AD and nondemented participants obtained on the Similarities subtest. As can be seen from the figure, consistent with the prediction that AD patients would have greater difficulty reasoning about the abstract relationships between words and concepts because of semantic memory retrieval deficits, AD patients were found to perform more poorly than nondemented participants on the WAIS-III Similarities subtest ($d = 1.3$; $t(92) = -5.9$, $p < .001$).

Nonverbal Abstract Thinking and Reasoning Results

To test the prediction that nonverbal analogical reasoning is not affected in early Alzheimer's dementia, the mean scores of the AD participants were compared to those of a matched sample of nondemented participants on two measures of nonverbal abstract thinking: (a) WAIS-III Matrix Reasoning subtest; and, (b) Raven's Coloured Progressive Matrices Test. The mean scores obtained by the AD and nondemented participants on both of these measures of nonverbal reasoning were not expected to differ. In contrast to verbal analogical reasoning tasks in which retrieval of conceptual information from semantic memory is required, these nonverbal reasoning measures were considered more "pure" tasks of reasoning and abstract thinking not involving memory retrieval. If reasoning abilities per se are not affected in early Alzheimer's Disease, then group performance differences on the WAIS-III Matrix Reasoning test and the RCPM should not be observed.

Table 12 shows the mean scores, standard deviations, and range of scores obtained by each group of participants on both the WAIS-III Matrix Reasoning subtest and the RCPM test.

WAIS-III Matrix Reasoning subtest analyses. For the Matrix Reasoning subtest, the mean age-scaled scores of the Alzheimer's group were compared to the mean age-scaled scores of the nondemented group using planned comparison t-tests. Contrary to predictions, AD patients in the early stages of the disease performed significantly worse than a matched sample of nondemented participants on the WAIS-III Matrix Reasoning subtest (t (92) = -4.68, p <.001). The effect size was large (i.e., d = .94) according to Cohen's criteria (Cohen, 1992) indicating that a large percentage of the nondemented controls were performing above the average person in the AD group.

Raven's Coloured Progressive Matrices analyses. For the RCPM, the mean percentile score of the AD group was compared to the mean percentile score of the nondemented group using planned comparison t-tests for independent samples. As can be seen from Table 12, the results of this analysis indicated that AD participants performed worse that nondemented participants on the RCPM (d = .84; t (92) = -4.49, p <.001). Consistent with the previous analyses, these results also indicate that the nonverbal reasoning abilities of AD patients in the early stages of the disease are negatively affected.

In order to conduct a more in depth analysis of the performance of both the cognitively intact sample and the AD patients on the Raven's, the RCPM items were reclassified according to the problem groupings established by Villardita (1985). Set I included 11 items (A_1, A_2, A_3, A_4, A_5, A_6, Ab_1, Ab_2, Ab_3, B_1, B_2) contingent on use of the principle of sameness or simple visuo-perceptual ability. Set II included 19 items (A_7, A_8, A_9, A_{10}, A_{11}, A_{12}, Ab_4, Ab_5, Ab_6, Ab_7, Ab_8, Ab_9, Ab_{10}, Ab_{11}, B_3, B_4, B_5, B_6, B_7) contingent on the use of the principle of symmetry. This included the analysis of the

length and direction of lines, the arrangement of dots, and the wideness of angles. Set III included 6 problems (Ab_{12}, B_8, B_9, B_{10}, B_{11}, B_{12}) contingent on the discovery of analogical relationships. Because the last set of problems specifically involved nonverbal analogical reasoning, the performance of both groups of study participants on Set III problems was of particular interest. AD participants were expected to perform just as well as nondemented participants on these analogies only if reasoning abilities remain intact in early Alzheimer's Disease.

For each participant, the number of correct problems solved in each set was computed. Figure 6 shows the mean RCPM set scores for the nondemented group and the Alzheimer's group. Analyses revealed that AD participants performed just as well as the nondemented group on Set I problems - problems involving simple matching and pattern completion (t (92) = - 1.85; p = .067). However, AD patients performed significantly worse than the matched sample of nondemented participants on Set II problems (d = 1.4; t (92) = - 5.33; p = .001) and Set III problems (d = .92; t (92) = - 5.12; p <.001). Thus, AD patients had difficulty solving analogical reasoning problems even when the terms of the analogy did not require retrieval of conceptual information from semantic memory. This finding suggests that abstract thinking and reasoning abilities per se are negatively effected early in the course of Alzheimer's Disease.

Error Analyses

GAT error analyses. The number of associative errors, categorical errors, appearance match errors, and miscellaneous errors made on the word and picture analogies was counted separately for each participant. To assist in the calculation of error counts, a VB script computer scoring program that summed different error types in

a Microsoft Access (Version 7) database was written. This computer scoring program also provided GAT total scores for each participant.

Error frequencies then were calculated by dividing the number of AD (or nondemented) participants whose most common error type was an associative error (or category/ appearance match/miscellaneous error) by the total number of participants within each group (i.e., $\underline{n} = 47$). The number of participants who did not make one particular type of error most commonly also was calculated. Thus, the percentage of participants who had error frequency ties (i.e., committed equal amounts of different errors) also was determined. Table 13 shows the percentages of AD and nondemented controls making different types of errors most commonly on GAT picture and word problems.

Two questions were of interest: 1. What is the most common type of error made by the Alzheimer's group on both word and picture problems? 2. Does the most common type of error made by the Alzheimer's group differ from that made by the nondemented group? To answer these questions, chi-square analyses were conducted. Overall, the results revealed that for both word and pictorial analogies respectively, the percentages of AD patients making associative, category, appearance match, and miscellaneous errors differed (X^2 [3, $\underline{n} = 47$] = 35.979, $\underline{p} < .001$; X^2 [3, $\underline{n} = 47$] = 23.723, $\underline{p} < .001$;). More specifically, the trend was for the greatest percentage of AD patients to make associate errors regardless of whether or not the problems were words or pictures.

Chi-square tests also were used to determine if equal proportions of AD and nondemented participants made associative, category, appearance match, and

miscellaneous errors. The results revealed that the percentages of AD and nondemented participants making each error type most commonly did not differ (X^2 [4, $\underline{n}$ = 47] = 8.486, $\underline{p}$ = .075; X^2 [3, $\underline{n}$ = 47] = 3.486, $\underline{p}$ =.323). Associative errors appeared to be the most common type of error made by both AD and nondemented participants.

Raven's Coloured Progressive Matrices error analyses. All of the problems on the RCPM were presented in the form "A is to B as C is to ?" (i.e., A:B::C:D). Thus, each problem had four terms: the A term; B term; C term; and, D term. The task was to select the relevant D term from among six answer choice options. For a subset of RCPM items, one erroneous answer choice option always was a geometrical appearance match to the C term of the problem (i.e., C-match error) and one erroneous answer choice option always was a geometrical appearance match to the B term (i.e., B-match error). The set of items in which this was the case included the following 15 problems: Ab_4, Ab_5, Ab_6, Ab_7, Ab_9, Ab_{12}, B_4, B_5, B_6, B_7, B_8, B_9, B_{10}, B_{11}, and B_{12}. Of these problems, six questions called for analogical reasoning (Villardita, 1985). The errors made on this latter subset of problems were of particular interest in this study.

Each participant 's performance on this particular subset of RCPM items was examined in detail. The number of times B-match and C-match errors were made when solving nonverbal analogies was computed for each participant. Error frequencies then were calculated by dividing the number of AD (or nondemented) participants whose most common error type was a B-match (or C-match) error by the total number of people within each group (i.e., $\underline{n}$ = 47). The percentages of participants who made an equal number of B-match and C-match errors also were counted. Table 14 shows the percentages of AD and nondemented controls making B-match and C-match RCPM

errors as well as the number of participants who made neither of these two error types most commonly (i.e., "tie" between B-match and C-match errors).

In terms of the study predictions, there was no evidence to suggest that the AD group would more frequently make one particular type of error. If B-match errors are most frequent, the interpretation might be that AD patients are exclusively attending to the B term of the problem. If, on the other hand, C-match errors are most frequent, this might suggest that the perceptual attributes of the C term were the most salient aspects of the problem. Furthermore, in this latter case, it is possible that the A and B terms have not even been encoded or that their relevance for correct problem solving has been forgotten.

To determine whether or not any one particular error type was most frequent for the Alzheimer's patients, a chi-square analysis was conducted. Overall, the percentages of AD patients making B-match errors, C-match errors, or ties differed (X^2 [2, $n = 47$] = 6.681, $p = .035$). C-match errors were more common than B-match errors (t (46) = -2.88; $p = .006$) for the AD group.

To determine whether or not there were any qualitative differences in errorful performance between AD patients and nondemented older adults, another chi-square analysis was conducted. The most common error made did not vary as a function of group (X^2 [2, $n = 94$] = 4.907, $p = .09$). That is, equal proportions of AD and nondemented participants made RCPM B-match or C-match errors, or chose no one particular error type most frequently.

In summary, the results of the RCPM error analyses showed that AD patients tended to choose answer choice options that matched the perceptual attributes of the C

term of the analogy most frequently, perhaps because they were having difficulty attending to the other terms of the analogy. Furthermore, the performance of the AD patients could not be distinguished from that of the nondemented participants on the basis of the RCPM error analyses. In particular, the differences in the proportions of AD patients versus nondemented participants who made B-match errors, C-match errors, or "ties" were not significant.

Discussion

In this study, the abstract thinking and reasoning performance of nondemented older adults and Alzheimer's patients with mild dementia was examined using both verbal and nonverbal reasoning measures. Overall, the results indicated that even in the early stages of Alzheimer's Disease, reasoning abilities are compromised. That is, AD patients performed significantly more poorly on both verbal (i.e., WAIS-III Similarities subtest; Goranson's Analogy Test) and nonverbal (i.e., Raven's Coloured Progressive Matrices Test; WAIS-III Matrix Reasoning subtest) abstract thinking and reasoning measures compared to a sample of nondemented older adults matched on age, education, and gender variables.

Although these study findings suggest that abstract thinking and reasoning abilities deteriorate in early Alzheimer's Disease, one limitation of the study is that a relatively small sample of AD participants was tested. Furthermore, in the absence of random sampling, it is possible that the AD patients that were tested in this study were not representatitive of the population of AD patients found in other Canadian cities. Thus, one is cautioned not to generalize the results of this study to other populations of AD patients outside of the Greater Victoria, Comox Valley, and Lower Mainland areas

until further replications of this study are conducted with other more representative samples.

Even though the results of this study suggest that on average, AD patients showed reduced abstract reasoning performance, it also should be noted that some individuals with Alzheimer's Disease performed very well on tests of verbal and nonverbal reasoning. That is, for some AD patients, abstract thinking and reasoning abilities did not decline. Future research might be warranted to determine what premorbid factors predict that reasoning abilities will remain intact with the onset of a neurodegenerative disease like Alzheimer's dementia.

The findings from this study suggest that abstract thinking and reasoning deficits can be considered valid inclusion criteria on diagnostic sets for Alzheimer's dementia such as the DSM-IV (American Psychiatric Association, 1994) and the NINCDS-ADRDA criteria (McKhann et al., 1984). In their present form, both of these criteria sets stipulate that deficits in two or more areas of cognition must be observed in order for probable Alzheimer's Disease to be diagnosed.

Despite continued debate about what tests to administer to measure specific cognitive abilities, both diagnostic criteria sets acknowledge that certain functional areas should be assessed. In the case of the DSM-IV, abstracting deficits are specifically listed and in the case of the NINCDS-ADRDA criteria, it is recommended that problem-solving measures be administered because this area of cognitive functioning may potentially be impaired. However, in the absence of an extensive empirical foundation, the inclusion of abstract thinking and problem solving deficits as a potential marker for Alzheimer's Disease was questionable. Furthermore, after reviewing the literature, there was a lack of

evidence (or conflicting evidence) for whether or not reasoning even declines in early Alzheimer's Disease.

Although some previous research had been conducted examining the abstract thinking and reasoning performance of AD patients, many of the studies reviewed failed to control for disease severity. Therefore, the performance of Alzheimer's patients in the early stages of the disease process may have been obscured by the inclusion of moderate to severely demented patients. Moderate to severely demented patients would be expected to have reasoning impairments because of the extent and degree of cortical atrophy and the increased number of neurofibrillary tangles and senile plaques throughout the entire cerebral cortex. On the other hand, there is no "common-sense" prediction that can be made regarding the reasoning performance of Alzheimer's patients early in the course of the disease.

Since Alzheimer's Disease severity was controlled in this study, it can now be concluded with more certainty that even in the early stages of the disease, reasoning abilities are affected. Knowing the approximate time course for the development of reasoning difficulties is important to: (a) accurately assess and diagnose the condition; (b) guide neuropsychologists in their choice of tests to administer during a diagnostic assessment; (c) identify symptoms and follow their progression as the disease progresses; (d) help caregivers develop realistic expectations of the patient's abilities; and, (e) assist caregivers with utilizing appropriate communication strategies.

In terms of assisting caregivers, knowing that Alzheimer's patients with mild dementia have difficulties understanding analogies or thinking abstractly might impact the types of communication strategies that are used. For example, clinical practitioners

might recommend that information be presented in a clear and concise manner using concrete instructions. Furthermore, use of analogies in everyday conversation would not be recommended when attempting to communicate with Alzheimer's patients who would be expected to have difficulties understanding such comparisons. If caregivers were informed that Alzheimer's patients have difficulty reasoning and thinking abstractly, then they might not develop unrealistic expectations regarding what the patient should still be able to understand in the early stages of the disease process.

In terms of aiding neuropsychologists, these study findings offer justification for the inclusion of reasoning measures on test batteries designed for the purpose of dementia diagnosis. Clearly, by including abstract thinking and reasoning measures, the neuropsychologist is better able to examine the pattern of performance and make sense of the data. Knowing that there is empirical evidence to suggest that reasoning declines in early Alzheimer's dementia helps guide the neuropsychologist's diagnostic decision making process. Furthermore, in the event of repeat neuropsychological testing, the clinical neuropsychologist would expect reasoning impairments and perhaps further worsening of reasoning performance as the disease progresses.

Although the results of this study suggest that one important domain of thinking to assess when conducting an Alzheimer's dementia assessment is abstract thinking and reasoning, the study findings do not allow for any conclusions to be made with respect to the reasoning performances of individuals facing other neurodegenerative diseases. Future research should compare AD patients to patients with mild frontotemporal dementia on various measures of abstract thinking and reasoning to determine whether or not there are any quantitative or qualitative differences in performance. Such a study

may clarify how the data is to be interpreted by clinical neuropsychologists so that differential diagnoses can be made with more confidence. Currently, researchers at UCLA are using a computerized version of the GAT to study the verbal analogical reasoning performance of patients with frontotemporal lobe dementia.

While frontotemporal dementia patients are the obvious comparison group of choice because of predominant executive functioning impairments and hence possible abstract thinking and reasoning deficits, patients with other forms of dementia also should be investigated with respect to their reasoning abilities. Furthermore, more research is needed to examine the effects of both normal aging and dementia on other types of reasoning abilities (e.g., logical reasoning, sequential reasoning, social reasoning, conceptual reasoning, deductive reasoning). Such studies may offer clarification regarding whether specific reasoning abilities are differentially affected in early Alzheimer's dementia. On the other hand, AD patients may perform poorly on all tests of abstract thinking and reasoning because of extensive pathology in the prefrontal cortex, an area of the brain thought to be responsible for relational reasoning (Christoff, & Gabrieli, 2000; Fuster, 2000; Goel, & Dolan, 2000; Waltz et al., 1999).

Conclusions Regarding the Goranson Analogy Test

While a variety of abstract thinking measures assessing different aspects of reasoning exist, few tests of analogical reasoning have been developed that are appropriate for use with older adults. Thus, one unique aspect of this study included the creation and psychometric analyses of a new measure of analogical reasoning, the Goranson Analogy Test (GAT). Overall, the results of the psychometrical analyses of the GAT suggest that both versions were adequately reliable measures of verbal analogical

reasoning for use with older adults. For GAT (Version 1) and GAT (Version 2), Cronbach's alpha statistics were above .80 indicating a high degree of internal consistency. The results of two separate principle component analyses revealed adequate item unidimensionality for both 38-item GAT versions. In addition, using split-half procedures, the reliability estimations for GAT (Versions 1 and 2) were acceptable.

In terms of the validity of the GAT for measuring abstract thinking and reasoning, a review of the convergent and divergent validity correlations indicated that the GAT was positively correlated with other measures of abstract thinking and reasoning but not correlated with a measure of depression. Correlations generally falling within the .50 to .80 range existed between the GAT and other reasoning measures. Furthermore, the intercorrelations among other reasoning measures were no higher than those with the GAT, strengthening the case that this newly developed tests performs similarly in comparison to other tests of abstract thinking and reasoning. On the basis of these results, the GAT was judged to have adequate construct validity and can be considered a new test of reasoning ability.

The predictive validity of this new analogical reasoning measure also was assessed by conducting sensitivity and specificity analyses, by examining the positive predictive power of both GAT versions, and by examining the likelihood ratios associated with each version of the GAT. Overall, both versions of the GAT were able to disciminate among AD patients and nondemented older adults relatively well. Optimal sensitivity and specificity values were obtained a cut score of 27/38 for GAT (Version 1) and at a cut score of 28/38 for GAT (Version 2).

The clinical implications of these findings are clear. A score below these optimal

cut-scores should be accepted as indicating the presence of abstract thinking and reasoning deficits and are consistent with the presence of Alzheimer's Disease. In fact, examination of the positive predictive values of the GAT indicates that of all those participants scoring below the cut-off values, a large percentage had been previously diagnosed with Alzheimer's Disease. Furthermore, the results of the likelihood ratio analyses indicated that it is two (GAT Version 2) to six (GAT Version 1) times more likely to see scores below the cut off values in someone with AD than in someone without AD.

In comparison to the other reasoning measures administered in this study, the GAT performed similarly if not better. For example, the GAT was equally if not more sensitive and specific in terms of being able to differentiate between groups. Both GAT versions had comparable if not higher positive predictive power than the other reasoning tests. These newly developed measures also were able to improve the pre-test probability of detecting AD by a small (GAT Version 2) to moderately high (GAT Version 1) amount. Furthermore, the GAT generally increased the pre-test probability of detecting AD to a greater extent than the other reasoning measures.

While the purpose of this study was not to compare and contrast the psychometric properties of the two GAT versions, a preliminary examination of the data suggests that GAT (Version 1) may have better predictive validity. Overall, the sensitivity and specificity values, the positive predictive power, and the likelihood ratio associated with GAT (Version 1) were better than those associated with GAT (Version 2). The reasons for these differences remain unclear at the present time and warrant further study.

In summary, the results of all of the psychometric analyses conducted for the purpose of this study indicated that the 38-item GAT versions had acceptable reliability for measuring reasoning performance, acceptable internal consistency, and adequate unidimensionality. Preliminary data on the construct validity and predictive validity of the GAT also revealed promising results. Despite the psychometric soundness of this new test of reasoning ability, it should be noted that the GAT is to be used as a diagnostic adjunct in combination with other tests tapping different areas of cognitive functioning. Obviously, it should not serve as the sole criterion for diagnosing dementia of the Alzheimer's type. Furthermore, it remains to be seen whether or not some combination of reasoning tests, rather than the GAT alone, may be able to best differentiate cognitively intact from demented individuals.

Future research regarding the reliability and validity of the GAT also is needed. For example, an examination of test-retest reliability and alternate forms reliability should be undertaken considering that two versions of the GAT were created. In addition, the construct validity of the GAT could be examined more thoroughly using Rasch item response theory analysis (Fox, & Jones, 1998).

Rasch analysis is based on a mathematical model that takes into account the relationship between a given item's level of difficulty and a person's ability. In particular, the Rasch model provides the conditional probability of a binary outcome (i.e., correct/incorrect) given the person's ability level and the item difficulty level (Fox, & Jones, 1998). The expectation is that more people are able to answer easier items than harder items. Likewise, less skilled people should obtain fewer correct items than more skilled people. To determine whether the data conform to these model expectations, the

Rasch analysis provides fit statistics for each person and each item (Wright, & Stone, 1979 in Fox, & Jones, 1998). For example, when high ability people are unable to answer "easy" questions, poor "item fit statistics" are obtained and this provides a rationale for discarding the item.

There are many benefits for conducting Rasch analyses on the newly developed Goranson's Analogy Test. First, Rasch analyses would offer further validation of the GAT as a test that measures low and high degrees of verbal analogical reasoning ability. Thus, Rasch analysis would provide another indication of the construct validity of the GAT. Good item fit statistics would imply that items refer to a unidimensional ability and would indicate that various items measure different levels of verbal analogical reasoning. Alternatively, the GAT could be further refined in the event that poor item fit statistics were obtained.

Second, hierarchical placement of test items along a continuum would give an indication of whether test questions progress from easy to difficult. Thus, a determination of whether the existing GAT contains items of greater or easier difficulty than desired can be made.

Third, Rasch analyses would allow for yet another determination of the reliability of the GAT. For example, Rasch methodology could be used to obtain a person reliability estimate and an item reliability estimate. Person reliability indicates the degree to which items measure people in a consistent manner. Item reliability indicates the degree to which items relate to each other in a consistent way.

Rasch analysis could also provide another way to look at how well the GAT measures verbal analogical reasoning by estimating how well the test items are distinct

from each other and distinguish among people. In Rasch analysis, indicators of "Separation" provide this estimate. Furthermore, the Rasch model could provide a means of converting GAT raw scores to an interval-equivalent scale that could then be used to compare the GAT test versions more comprehensively.

In summary, Rasch analysis of the GAT could provide a measure of item difficulty, scale unidimensionality, construct validity, test reliability, item separation, and linearity. As such, more evidence for the usefulness of the GAT as a measure of analogical reasoning could be obtained.

Performance of Alzheimer's Patients on the Goranson Analogy Test

In this study, the GAT was administered to AD patients and a matched sample of nondemented older adults to determine: (a) if verbal analogical reasoning differs as a function of group membership; (b) if AD patients reason better with pictures compared to words. AD patients were predicted to have difficulty solving verbal analogies because of semantic memory retrieval deficits associated with the onset of Alzheimer's disease. Thus, the performance of AD patients on the GAT was expected to be worse than that of a matched sample of nondemented older adults. Furthermore, there was some limited evidence to suggest that AD patients might reason better when presented with pictorial analogies compared to word analogies.

Consistent with study expectations, AD patients performed more poorly than nondemented controls on both versions of the GAT. If verbal analogical reasoning depends on retrieval of conceptual information from semantic memory, this finding might imply that AD patients had more difficulties retrieving information from semantic

memory compared to nondemented controls. While this explanation is certainly plausible, there are other explanations that may also account for the study findings.

AD patients, for example, may not have encoded all of the terms of the analogy because of attention/concentration deficits. In this case, access to an intact semantic memory might never have occurred. According to this explanation, focal attentional deficits may have interfered with analogical reasoning performance. Alternatively, semantic memory accessing problems may have occurred in the absence of attention and concentration difficulties. An accessing deficit hypothesis could account for the study findings if one adopts a disconnection model similar to that of Farah and McClelland (1991, 1992). According to the PDP model of semantic memory proposed by Farah and McClelland, if routes of entry between the name/pictorial input units and functional semantic memory were partially degraded or disconnected, analogical reasoning performance would be negatively affected.

A third possibility is that the working memory abilities of Alzheimer's patients might have been impaired which in turn may have impacted analogical reasoning performance. Working memory is a memory store where individuals are temporarily able to hold in mind a main goal while simultaneously performing concurrent subgoals (Koechlin, Basso, Pietrini, Panzer, & Grafman, 1999). It is well known that working memory is an important prerequisite of many cognitive abilities such as the ability to engage in problem solving and reasoning whereby mental operations must be performed simultaneously and efficiently (Johnson-Laird, & Byrne, 1991; Keane, Ledgeway, & Duff, 1994).

According to the working memory deficit hypothesis, analogical reasoning performance deficits might occur because of difficulties maintaining analogical inferences in a limited capacity store of brief duration. More specifically, once a relational inference has been made, it has to be maintained in working memory while all of the answer choice options are processed and evaluated. If information can only be maintained for a brief duration or if information processing resources are at a premium, it might not be possible to simultaneously work on solving sub-goals while keeping in mind key inferences. Consequently, when working memory is overloaded, key inferences may be lost or forgotten and errorful analogical reasoning performance would occur. Thus, as an alternative to the semantic memory accessing/retrieval deficit hypothesis, AD patients may have performed more poorly than nondemented controls on the GAT analogies because of working memory inefficiencies.

Although the results of this study indicated that AD patients had more analogical reasoning difficulties compared to nondemented older adults, future research should compare and contrast the various explanations to account for these group differences. Although several hypotheses that can account for the study findings have been postulated, it remains to be determined whether the semantic memory retrieval deficit hypothesis is more effective than the encoding deficit hypothesis or the working memory deficit hypothesis in accounting for group performance differences.

In terms of whether modality of presentation affects the analogical reasoning performance of AD patients, the findings from this study did not indicate that reasoning with pictures was easier than reasoning with words for AD patients, most of whom were recruited from the Greater Victoria area. This is in contrast to the study prediction of a

pictorial advantage and is generally not consistent with the results of previous research (i.e., Bub et al., 1988; Chertkow et al., 1988). On the other hand, previous researchers asked demented patients to answer questions when names and pictures of both living and nonliving things were presented. The living/nonliving distinction was not a factor that was taken into consideration in this study when creating the analogical reasoning problems for inclusion on the GAT.

It is possible that the organic status of the terms of each analogy was a significant but uncontrolled confounding factor that could have impacted the extent to which modality of presentation influences analogical reasoning performance. In this study, however, modality of presentation was not a significant factor impacting analogical reasoning performance. On GAT (Version 1), for example, AD patients were equally able to solve picture and word analogies. For GAT (Version 2), weak evidence for a word advantage was observed, but the effect size was small indicating that statistical power associated with the test result was low.

The fact that a word advantage was not consistently observed across GAT versions does not support the view that modality of presentation affects analogical reasoning performance in this study context. On the other hand, the findings can not be solely attributed to a problem content confound considering the design of the GAT tests. Specifically, the GAT tests were created such that problem content was controlled for across versions while orthogonally manipulating modality of presentation. The problem content of the word analogies on one version, for example, matched the problem content of the picture analogies on the other version. Thus, the relationships governing the analogies remained constant across word and picture problems.

If modality of presentation had affected analogical reasoning performance, the same modality advantage would have consistently been observed across both GAT versions. This finding would have implied that the problem content that the examinee previously found easy on one version of the GAT suddenly became more challenging when presented in the opposite modality on the second version of the GAT. Alternatively, if problem content had been the primary factor influencing analogical reasoning performance, the opposite modality advantage would have been observed across GAT versions. This finding would have implied that the problem content that the examinee found easy on one version of the test was also easy on the second version of the test when presented in a different modality.

Overall, the results of this study suggested that the modality of presentation did not significantly and consistently impact analogical reasoning performance. That is, reasoning with words was generally equally difficult compared to reasoning with pictures for both the AD group and the nondemented control group who participated in this study. Because the same modality advantage was not observed across GAT versions, the modality hypothesis was not supported. However, by controlling for problem content, this study is unique in that it allows one to conclude with more certainty that modality of presentation does not influence reasoning performance. Failure to observe the opposite modality advantage across GAT versions also does not support the view that variations in problem content difficulty was a significant factor affecting interpretation of the study results.

There are, however, some study limitations to consider. First, the same study participants did not take both versions of the GAT. For example, the performance of an

individual completing GAT (Version 1) word problems was compared to the performance of another individual completing GAT (Version 2) picture problems. The drawback to using this type of research design is that between subject error variance might account for any observed modality differences. However, there were no demographic differences between the group of participants who were administered GAT (Version 1) and the group of participants who were administered GAT (Version 2). Regardless, the two groups may have differed on some variable not considered by the researcher. Ideally, both GAT versions should be administered to the same sample of participants in the future to confirm the absence of modality findings.

Another limitation of this study concerns the representativeness of the sample of participants who were administered the GAT. The group of AD patients and nondemented participants who participated in this study were relatively well educated. In a well educated sample, it is possible that modality of presentation is not a strong factor affecting analogical reasoning performance. In a less educated sample, on the other hand, it is possible that modality of presentation could more strongly influence analogical reasoning performance. In particular, if word vocabulary knowledge is lacking, analogical reasoning performance might not be as efficient when the terms of the analogy are presented in word format compared to when they are presented in pictorial format. Indeed, previous research with children has shown that if the contextual knowledge underlying the terms of the analogy is absent, children are unable to formulate correct analogical reasoning responses (Goswami, & Brown, 1989).

Future research also should examine whether or not modality of presentation fails to impact performance only when certain types of analogical relationships are presented. For

example, the results of this study do not allow one to determine whether or not product-producer analogies (e.g., "Milk is to cow as juice is to orange") are more susceptible to modality effects compared to functional category relationships (e.g., "Airplane is to bird as boat is to fish"). Previous research has identified a large number of analogical relationships, many of which have been incorporated into the design of the GAT problems. However, the relationships are easier to understand in word versus pictorial format. Taken together, the results of this study merely suggest that modality of presentation is not a significant factor impacting the ability to solve many different kinds of analogy problems.

Future researchers who are interested in using the GAT to examine modality effects on reasoning performance also should strive to equate the number of study participants receiving both GAT versions. Because an attempt was made to control for severity of dementia in this study, two AD patient's data were dropped from the final analyses. Unfortunately, this resulted in unequal numbers of AD patients receiving each version of the GAT. Statistically, the impact of having unequal numbers of participants receive GAT (Version 1) compared to GAT (Version 2) is not too problematic unless group sizes are sharply unequal which will result in violation of homogeneity of covariance. It is therefore recommended that future studies strive to equate the number of participants who receive GAT Versions 1 & 2 so that the interpretation of any observed modality effects is clear.

Implications for the PDP Model of Semantic Memory

It has been proposed that Farah and McClelland's Parallel Distributed Processing Model of Semantic Memory (1991, 1992) can be used to explain the verbal analogical reasoning performance of Alzheimer's patients if one assumes that the retrieval of conceptual information is a necessary component of analogical reasoning performance.

In this study, AD patients performed worse than nondemented controls on the GAT analogical reasoning problems. According to the PDP model, this finding can be explained by hypothesizing that there are degraded connections between word and picture units and functional semantic memory due to the onset of Alzheimer's dementia preventing correct retrieval of conceptual information from semantic memory.

There are several drawbacks to using this model to account for the study findings. First, the PDP model of semantic memory is a "disconnection" model that can be used to explain reasoning deficits in Alzheimer's disease by assuming that the functional semantic memory units become disconnected from either the word units or the pictorial units. The question remains as to whether analogical reasoning difficulties are observed because of problems "accessing" functional semantic memory or "retrieving" information from functional semantic memory.

Despite a rather large body of literature showing that some Alzheimer's patients have no difficulties accessing semantic memory, it is possible that analogical reasoning requires sustained attention so that all terms of the analogy access semantic memory quickly and in parallel. If the capacity for controlled attention declines with the onset of Alzheimer's disease, it is possible that the terms of the analogy were not encoded properly and so were not able to access semantic memory. On the other hand, even if attentional mechanisms remain functional in early AD, it is possible that partially degraded connections between word and pictorial units and functional semantic memory caused insufficient activation of the relevant conceptual nodes and/or inappropriate activation of irrelevant associations in functional semantic memory. Still another

possibility is that partially degraded connections interfered with the retrieval of conceptual information from functional semantic memory.

To summarize, the analogical reasoning performance of AD patients could have been worse than that of nondemented controls because of degraded connections preventing efficient semantic memory access *and* because of degraded connections preventing efficient semantic memory retrieval or because of attention/concentration deficits. By using Farah and McClelland's PDP model of semantic memory to account for analogical reasoning deficits, one is not able to determine whether or not the underlying difficulty related more to problems accessing functional semantic memory or difficulties retrieving conceptual relationships from functional semantic memory.

Furthermore, although semantic memory plays an important role in analogical reasoning, other aspects of cognition are important to consider. For example, it is possible that the capacity for abstract reasoning may be related to inhibitory failures. That is, AD patients may have had difficulty inhibiting highly-related associational information activated in functional semantic memory and therefore may have chosen incorrect answer choice options. Indeed, preliminary research suggests that patients with frontotemporal dementia have difficulty solving verbal analogies because of difficulties inhibiting irrelevant information that is semantically or perceptually similar to the analogy terms (Morrison et al., 2000).

While Farah and McClelland's PDP model of semantic memory provides a useful framework for understanding the analogical reasoning performance of AD patients in comparison to nondemented controls, this disconnection model may not be comprehensive enough to explain the complex task of reasoning by analogy. Obviously,

there is a need to develop a coherent neuropsychological model of reasoning by analogy especially when the problems are presented in the form, "A is to B as C is to ?".

Although Sternberg (1977) developed a cognitive model of analogical reasoning performance, the model is limited in that: (a) it is a modular model that fails to delineate the neuroanatomical substrates involved in analogical reasoning; (b) it does not adequately account for the role of attention/concentration, working memory, and inhibition mechanisms in analogical reasoning performance; and (c) it does not specifically account for any analogical reasoning performance deficits. The goal would be to develop a model of verbal analogical reasoning that takes into account all of these factors.

Other theories of analogy [i.e., Holyoak and Thagard's (1989) Analogical Constraint Mapping Engine, the Structure-Mapping Engine of Falkenhainer et al. (1989); Keane and Brayshaw's (1988 in Keane et al., 1994) Incremental Analogy Machine (1988 in Keane et al., 1994)] have been proposed, but have been criticized (i.e., Keane et al., 1994) for failing to account for the neurological mechanisms guiding analogical thought. Furthermore, these cognitive models were designed to solely model how "analogical mappings" are performed. That is, the goal has been to develop models explaining how people draw comparisons between concepts, what cognitive representations are made during this process, what algorithms are followed, and what constraints limit analogical mapping. Little attention has been paid to other subprocesses besides analogical mapping that may be involved in analogical thinking.

If analogical reasoning is subserved by an interconnected system of cortical regions, then the following neuropsychological model of analogical reasoning is

proposed. First, all of the terms of the analogy would have to be encoded into semantic memory. Effective encoding would require the ability to focus and sustain attention. Once encoded into semantic memory, conceptual information related to the terms of the analogy would need to be activated and retrieved from functional semantic memory, a process involving bilateral hippocampal activation. Having done this retrieval, one could then infer the relations between concepts by performing analogical mappings between the terms of the analogy. Specifically, one would have to infer how the A and B terms of the analogy are alike, and map this relationship on to the C term of the analogy in order to determine the correct D term that best completes the problem.

While various theories of analogical mapping have been proposed, all seem to imply that analogical mapping occurs in a parallel distributed processing memory network as the result of excitatory and inhibitory activation of interconnected nodes. According to Holyoak and Thagard's ACME model, for example, when the network settles into a stable state, the nodes whose activation exceeds a threshold corresponds to the optimal match.

Once the mapping stage is complete, the inferred relationship between the A and B terms would have to be maintained in working memory while evaluating all of the possible match options to the C term of the analogy. Since neuroimaging studies show that working memory is mediated by the dorsolateral prefrontal cortex (Prabhakaran, Narayanan, Zhao, & Gabrieli, 2000; Fuster, 2000), it is likely that this brain region would be involved in this stage of analogical problem solving.

Furthermore, in a recent review of neuroimaging studies of reasoning performance, Christoff and Gabrieli (2000) found that not only is the dorsolateral

prefrontal cortex involved during reasoning, the frontopolar areas are activated as well. While the dorsolateral prefrontal cortex is thought to be involved in maintaining information in working memory, the frontopolar cortex is thought to be involved in hypothesis generation and evaluation (Baker et al., 1996). Hypothesis generation and evaluation needs to be carried out when deciding which answer choice option best completes the analogy. In fact, Prabhakaran and colleagues (1997) observed frontopolar activation during performance of the Raven's Progressive Matrices task and argued that this cortical region is involved during the inductive reasoning process when hypotheses regarding which term best completes the analogy are being generated and evaluated.

Another cognitive information processing ability that is probably required to reason analogically includes the ability to inhibit irrelevant information that is activated in memory but that does not conform to the constraints of the problem. For example, when answer choice options that are highly related to the C term of the analogy are presented, it is necessary to be able to inhibit these erroneous answer choices from interfering with correct problem solution. Since inhibitory mechanisms have been shown to be mediated by frontal lobe mechanisms, it is understandable that activation of the frontal lobes would be detected during neuroimaging studies of reasoning performance.

In summary, verbal analogical reasoning requires intact attention/concentration, intact access to and retrieval from functional semantic memory, intact working memory, intact inhibitory mechanisms, and intact inductive reasoning and abstract thinking abilities involving "mapping" of relationships. Furthermore, functional magnetic resonance imaging studies (e.g., Goel, & Dolan, 2000; Prabhakaran et al., 1997) are beginning to outline the neural substrates of inductive inference making, and some have

even used analogical reasoning tests to study the brain mechanisms involved. Early evidence suggests that bilateral hippocampal activation, frontopolar activation, and dorsolateral prefrontal cortex activation occurs during analogical reasoning.

In contrast to verbal analogical reasoning, when analogical reasoning needs to be carried out in the nonverbal domain, intact visuospatial and visuoperceptual skills are necessary. Thus, parietal lobe involvement is likely. While activation of this region has not been reported in the limited number of functional imaging studies employing reasoning tasks, lesion studies (e.g., Berker, & Smith, 1988; Costa, 1976) reveal that patients with posterior parietal lobe deficits often demonstrate impaired performance on the RCPM.

Obviously, more research is needed to refine this preliminary neurocognitive model of analogical reasoning. By establishing a better understanding of the neural mechanisms as well as the cognitive information processing abilities involved in both verbal and nonverbal analogical reasoning, we may be better able to develop a more comprehensive understanding of why analogical reasoning deteriorates with the onset of Alzheimer's Disease.

Nonverbal Reasoning Performance of Alzheimer's Patients

Nonverbal reasoning tasks are similar to verbal reasoning tasks in that both require attention/concentration, working memory, inhibitory mechanisms, and the evaluation of internally generated information in order to arrive at a logical answer. However, nonverbal reasoning tasks, unlike verbal reasoning tasks, also require intact visuoperceptual/ visuospatial abilities. As well, nonverbal reasoning tasks are less dependent on the retrieval

of conceptual information from semantic memory and therefore are presumed to offer a more "pure" measure of the reasoning ability of Alzheimer's patients.

On both of the nonverbal reasoning tasks that were administered as part of this study (i.e., Matrix Reasoning; RCPM), Alzheimer's patients performed more poorly than nondemented controls. This finding was not in keeping with expectations and suggests that even in the early stages of Alzheimer's disease, abstract thinking and reasoning abilities are negatively affected.

Furthermore, when the RCPM problems were reclassifed according to the problem groupings established by Villardita (1985), Alzheimer's patients performed more poorly than nondemented controls on the set of problems specifically measuring analogical reasoning performance. Thus, even when problems that only required visuoperceptual matching abilities were removed, the RCPM performance of the AD group was still deficient relative to the performance of nondemented controls.

Although the results of this study suggest that abstract thinking and reasoning abilities decline with the onset of Alzheimer's Disease, further research examining why performance is negatively affected will need to be conducted. Considering the variety of cognitive processes involved in analogical reasoning, it is possible that nonverbal reasoning performance declines for a multitude of reasons including: attention/concentration difficulties, working memory problems, visuospatial/visuoperceptual difficulties, inhibitory control mechanism failure, and/or inferential reasoning impairments including an inability to perform the correct relational mappings. Future studies should compare and contrast the contribution of each of these abilities when engaging in nonverbal reasoning in order to determine which ability most significantly interferes with correct performance.

Additionally, examining the nonverbal analogical reasoning performance of a larger sample of older adults across different age cohorts might be beneficial to establish whether or not there are any age effects. Furthermore, the question remains as to whether or not nonverbal analogical reasoning performance also declines with the onset of other dementing illnesses such as frontotemporal dementia or vascular dementia, for example.

Even though the results of this study show that the nonverbal reasoning performance of AD patients was worse than that of a matched sample of nondemented patients, the mean RCPM percentile score and the mean Matrix Reasoning age-scaled score of the AD group still fell within normal limits. That is, when the nonverbal reasoning scores of the AD group were interpreted in relation to the performance of the standardization sample used to develop the WAIS-III and RCPM norms, the AD patients in this study performed in the "average range". Clinically, this suggests that despite a significant decline in reasoning abilities, abstract thinking and reasoning abilities are still not technically "impaired" in early Alzheimer's Disease.

On the other hand, the finding of average range performance for the AD group may relate to the claim that the WAIS-III and RCPM norms are poor for the elderly. The WAIS-III norms, for example, have been criticized for offering lower ability estimates because of test administration error. For example, WAIS-III test developers did not take into account fatigue affects when administering long test batteries to older adults. The implication for this study is that the percentile scores obtained by the AD and nondemented participants may be falsely elevated.

Similarly, there are problems with the standardization sample on which the RCPM norms were based. In particular, the standardization sample was not drawn from a

representative sample of older adults screened for medical illness or cognitive status, but rather was drawn from a sample of older adults visiting a local medical clinic in Britain. Furthermore, these norms do not take into account the effects of education or gender, factors that have been found to impact RCPM performance (Measso et al., 1993). As a result, the RCPM norms may be too low, resulting in higher percentile scores for the sample of highly educated AD patients examined in this study.

Despite the limitations associated with the normative data used in this study, the results of this cross-sectional research design reveal that relative to a control sample of nondemented adults matched on age, education, and gender variables, clinically meaningful deterioration in nonverbal reasoning abilities has occurred for AD patients. In this study, large effect sizes were obtained indicating that the group differences in performance on the WAIS-III Matrix Reasoning subtest and the RCPM were indeed significant. Thus, it can be concluded that significant change in reasoning performance occurred as the result of the onset of Alzheimer's Disease.

It is possible that the sample of AD participants who took part in this study had superior reasoning abilities premorbidly and then evidenced a decline in reasoning performance as the result of the onset of Alzheimer's Disease. Alternatively, the AD samples' percentile scores on the various reasoning measures may have been falsely elevated because of inherent flaws in the normative databases used to derive the percentile scores. That is, instead of concluding that the AD group demonstrated "average range" performance, it is possible that with reference to a more representative standardized sample, they are really demonstrating "impaired to low-average" range performance. Regardless, the fact that statistically significant group differences and large effect sizes were observed

indicates that relative to a matched sample of nondemented controls, the nonverbal abstract thinking and reasoning abilities of AD patients declined with the onset of dementia.

Reliable Change

In this study, a cross sectional research design was used to show that meaningful change in abstract thinking and reasoning abilities occurred for AD patients early in the course of the disease. That is, the reasoning performance of AD patients was worse than that of a matched sample of nondemented controls which was interpreted to mean that deterioration in the premorbid reasoning abilities of AD patients had occurred.

Future research might focus on replicating the study findings using reliable change methodologies based on test-retest protocols rather than cross-sectional research designs. That is, it might be useful to determine how much change is to be expected at retest when cognitively intact individuals are re-administered various abstract thinking measures such as the WAIS-III Matrix Reasoning and Similarities subtests, the RCPM, and the GAT. Such information would be useful for geriatric neuropsychologists conducting repeat neuropsychological assessments, which is becoming a more common practice prior to establishing a dementia diagnosis.

When re-test data are worse than baseline data, clinical neuropsychologists must be careful in deducing that decline in abstract thinking and reasoning abilities has occurred without considering other factors that affect score change such as measurement error. For example, the reliability of the test must be considered as well as statistical effects like regression to the mean. Other variables such as practice effects and order effects and their interaction with factors such as decline due to normal aging must be considered. Given all of these considerations, it is important to determine what constitutes clinically reliable and

significant test score change.

There are many different statistical formulae that have been proposed for assessing reliable change at retest. Some of these methods (i.e., regression-based procedure of Crawford and Howell, 1998) do a better job in comparison to others in accounting for confounding variables and measurement error. Because the measurement of change over time using repeated assessments is of increasing concern to neuropsychologists, it has been recommended (e.g., Sawrie, Chelune, Naugle, & Luders, 1996) that future manuals for neuropsychological tests include reliable change methods as part of their normative process to aid neuropsychologists in adequately assessing cognitive change. Thus, future research related to the GAT and the other reasoning measures used in this study might focus on using regression-based reliable change score methodologies to demarcate the upper and lower bounds of confidence intervals. That is, reliable change methodologies should be employed to outline a range of scores suggestive of "normal" test score change attributable to measurement error. Then, if any scores fell outside of this confidence interval range on repeat testing, neuropsychologists would be able to more confidently conclude that "true" abstract thinking and reasoning changes have taken place.

Relevance of the Error Analyses

Two error analyses were conducted as part of this study. First, the GAT error analyses revealed that when AD patients are engaged in verbal analogical reasoning, they have difficulty inhibiting highly-related answer choice options that are strongly associated with the C term of the analogy. Indeed, the most common GAT error made by the AD group was an associative error while categorical errors and appearance match errors were less common. Associative errors also were the most common type of error made by the

nondemented group. Thus, the verbal analogical reasoning performance of the AD group and the nondemented group could not be distinguished on the basis of the GAT error analyses.

The finding that associative errors were the most common type of error made on the GAT by AD patients is in keeping with findings from other studies which have shown that dementia patients generally have difficulty inhibiting highly-related semantic information (Morrison et al., 2000). However, in contrast to previous research (i.e., Gentner, & Toupin, 1986 in Goswami, & Brown, 1989) with young children and novice learners, who tend to solve verbal analogies on the basis of mere appearance or shared object attributes, the findings from this study suggest that AD patients and nondemented older adults are unduly influenced by highly-related semantic information.

The second error analysis that was conducted as part of this study also revealed that there were no qualitative differences in errorful performance between AD patients and nondemented older adults when solving nonverbal analogies. For both groups, the most frequent type of error made on the Raven's Coloured Progressive Matrices Test was a C-match error. That is, participants tended to most frequently choose the erroneous answer choice option that was a geometrical appearance match to the C term of the analogy. Choosing a geometrical appearance match to the B term of the analogy was less common. These findings imply that the nonverbal reasoning performance of Alzheimer's patients and nondemented older adults could not discriminated on the basis of the types of errors made.

Both the GAT error analyses and the RCPM error analyses were conducted in an attempt to discover something about the cognitive processes underlying verbal and

nonverbal analogical reasoning performance. Overall, both the GAT error analyses and the RCPM error analyses revealed that errorful performance occurred because of a tendency to disregard underlying relational commonalities between all of the terms of the analogy. Instead, for both verbal and nonverbal analogical problem solving, study participants tended to focus their attention exclusively on the C term of the analogy and use strategies based on perceptual similarity (RCPM) or relational similarity (GAT) to solve the problems.

Despite the fact that the types of errors most frequently made by the AD patients did not differ from those of the nondemented controls, future research should examine whether or not the reasons for errorful performance differ between groups. Essentially, two questions remain, "Why do AD patients and nondemented older adults commit analogical reasoning errors?" and "Do the reasons for errorful performance differ for the AD and nondemented groups?". For example, it may be that the underlying cognitive information processing inefficiencies that result in errorful performance for the AD group are different than that of older adults who are aging normally. AD patients, for example, may not even encode the A and B terms of the analogy because of attention/concentration deficits whereas adults who are aging normally, may encode these terms but may commit errors because of working memory failures. Future research seems warranted to compare and contrast various theories accounting for errorful performance when both verbal and nonverbal analogies are presented and when both demented and cognitively intact older adults are tested.

Summary

The results of the experiment reported in this dissertation indicate that both verbal and

nonverbal abstract thinking and reasoning abilities decline with the onset of Alzheimer's Disease. AD patients performed worse than nondemented controls on a variety of reasoning measures, including a newly developed test of analogical reasoning, the Goranson Analogy Test.

Although the results of this study can be partially interpreted with reference to Farah and McClelland's PDP model of semantic memory, it has been argued that this model may too limiting to account for the complex task of reasoning by analogy. It has been proposed that the capacity for analogical reasoning not only depends on intact semantic memory functioning and intact memory retrieval abilities, but also on attention/concentration, working memory, inferential "mapping" ability, and the inhibitory control of interference. In addition to these cognitive processes, nonverbal reasoning performance also requires intact visuoperceptual and visuospatial abilities. Thus, reasoning is a higher-order thinking skill that requires the intact functioning of an integrated network of cognitive processing resources subserved by frontal-temporal-parietal feedback loops.

The study findings not only have implications for theories of abstract thinking and reasoning, they also have important implications for the nosological questions facing clinical researchers and practitioners interested in establishing diagnostic criteria sets for Alzheimer's dementia. Empirical support for the contention that abstract thinking and reasoning is deficient in mild Alzheimer's dementia validates the inclusion of these symptoms on current diagnostic criteria sets and hopefully helps clinical neuropsychologists in the diagnostic decision making process.

References

Albert (1988). Cognitive function. In M. Albert, & M. Moss (Eds.), Geriatric neuropsychology (pp. 33-53). New York: The Guilford Press.

Albert, M., Duffy, F., & Naeser, M. (1987). Nonlinear changes in cognition and their neurophysiologic correlates. Canadian Journal of Psychology, 41, 141-157.

American Psychiatric Association (1994). Diagnostic and Statistical Manual of Mental Disorders (4th ed.). Washington, DC: Author.

Anastasi, A., & Urbina, S. (1997). Psychological testing (7th ed.). Upper Saddle River, NJ: Prentice-Hall, Inc.

Anderson, J., Hartley, A., Bye, R., Harber, K., & White, O. (1986). Cognitive training using self-discovery methods. Educational-Gerontology, 12(2), 159-171.

Andree, B., Hittmair, M., & Benke, T. (1992). Recognition and explanation of proverbs in Alzheimer's disease. Journal of Clinical and Experimental Neuropsychology, 14, 372.

Axelrod, B., & Henry, R. (1992). Age-related performance on the Wisconsin Card Sorting, Similarities, and Controlled Oral Word Association Tests. Clinical Neuropsychologist, 6, 16-26.

Babcock, R. (1994). Analysis of adult age differences on the Raven's Advanced Progressive Matrices Test. Psychology and Aging, 9(2), 303-314.

Babcock, R., & Laguna, K. (1996). An examination of the adult age-related differences on the Raven's Advanced Progressive Matrices: A structural equations approach. Aging, Neuropsychology and Cognition, 3(3), 187-200.

Baker, S., Frith, C., Frackowiak, R., & Dolan, R. (1996). Active representation of shape and spatial location in man. Cerebral Cortex, 6, 612-619.

Barona, A., Reynolds, C., & Chastain, R. (1984). A demographically based index of premorbid intelligence for the WAIS-R. Journal of Consulting and Clinical Psychology, 52, 885-887.

Berker, E., & Smith, A. (1988). Diaschisis, site, time and other factors in Raven performances of adults with focal cerebral lesions. International Journal of Neuroscience, 38, 267-285.

Bieliauskas, L. (1993). Depressed or not depressed? That is the question. Journal of Clinical and Experimental Neuropsychology, 15(1), 119-134.

Blair, J. & Spreen, O. (1989). Predicting premorbid IQ: A revision of the National Adult Reading Test. The Clinical Neuropsychologist, 3(2), 129-136.

Bonilla, J., & Johnson, M. (1995). Semantic space in Alzheimer's disease patients. Neuropsychology, 9(3), 345-353.

Brown, A. (1989). Analogical learning and transfer: What develops? In S. Vosniadou , & A. Ortony (Eds.), Similarity and analogical reasoning (pp. 369-412). Cambridge University Press.

Brown, A., & Kane, M. (1988). Preschool children can learn to transfer: Learning to learn and learning from example. Cognitive Psychology, 20, 493-523.

Bub, D., Black, S., Hampson, E., & Kertesz, A. (1988). Semantic encoding of pictures and words: Some neuropsychological observations. Cognitive Neuropsychology, 5, 27-67.

Butters, N., Granholm, E., Salmon, D., Grant, I., & Wolfe, J. (1987). Episodic semantic memory: A comparison of amnestic and dementia patients. Journal of Clinical and Experimental Neuropsychology, 9, 479-497.

Canadian Study of Health and Aging Work Group (2001). Canadian Study of Health and Aging- 3: Neuropsychological Assessment Manual.

Caramazza, A., Hillis, A., & Rapp, B. (1990). The multiple semantics hypothesis: Multiple confusions? Cognitive Neuropsychology, 7, 161-189.

Carswell, L, Graves, R., Snow, W.G., & Tierney, M. (1997). Postdicting verbal IQ of elderly individuals. Journal of Clinical and Experimental Neuropsychology, 19(6), 914-921.

Chan, A., Butters, N., Salmon, D., & McGuire, K. (1993). Dimensionality and clustering in the semantic network of patients with Alzheimer's disease. Psychology and Aging, 8(3), 411-419.

Charter, R. (1996). Revisiting the standard errors of measurement, estimate, and prediction and their application to test scores. Perceptual and Motor Skills, 82, 1139-1144.

Chertkow, H., Bub, D., & Caplan, D. (1992). Constraining theories of semantic memory processing: Evidence from dementia. Cognitive Neuropsychology, 9(4), 327-365.

Chertkow, H., Bub, D., & Seidenberg, M. (1989). Priming and semantic memory loss in Alzheimer's disease. Brain and Language, 36, 420-446.

Christensen, K., Multhaup, K., Nordstrom, S., & Voss, K. (1991). A new cognitive battery for dementia: relative severity of deficits in Alzheimer's disease. Developmental Neuropsychology, 7(4), 435-449.

Christoff, K., & Gabrieli, J. (2000). The frontopolar cortex and human cognition: Evidence for a rostrocaudal hierarchical organization within the human prefrontal cortex. Psychobiology, 28(2), 168-186.

Chronbach, L. (1951). Coefficient alpha and the internal structure of tests. Psychometrika, 16, 297-334.

Clayton, V., & Overton, W. (1976). Concrete and formal operational thought processes in young adulthood and old age. International Journal of Aging and Human Development, 7, 237-246.

Cohen, G. (1981). Inferential reasoning in old age. Cognition, 9, 59-72.

Cohen, J. (1992). A power primer. Psychological Bulletin, 112(1), 155-159.

Cooke, D., & Michie, C. (1997). An item response theory analysis of the Hare Psychopathy Checklist - Revised. Psychological Assessment, 9(1), 3-14.

Costa, L.D. (1976). Interset variability on the Raven Coloured Progressive Matrices as an indicator of specific ability deficit in brain-lesioned patients. Cortex, 12, 31-40.

Cowart, C., & McCallum, S. (1984). Simultaneous-successive processing across the life-span: A cross-sectional examination of stability and proficiency. Experimental Aging Research, 10(4), 225-229.

Crawford, J., & Howell, D. (1998). Regression equations in clinical neuropsychology: An evaluation of statistical methods for comparing predicted and obtained scores. Journal of Clinical and Experimental Neuropsychology, 20(5), 755-762.

Daum, I., Riesch, G., Sartori, & Birbaumer, N. (1996). Semantic memory impairments in Alzheimer's disease. Journal of Clinical and Experimental Neuropsychology, 18(5), 648-665.

Diesfeldt, H. (1990). Recognition memory for words and faces in primary degenerative dementia of the Alzheimer's type and normal old age. Journal of Clinical and Experimental Neuropsychology, 12(6), 931-945.

Dijkers, M. (1997). Measuring the long-term outcomes of traumatic brain injury: A review of the Community Integration Questionnaire. Journal of Head Trauma Rehabilitation, 12(6), 74-91.

Falkenhainer, B., Forbus, K., & Gentner, D. (1989). Structure-mapping engine. Artificial Intelligence, 41, 1-63.

Farah, M., , & McClelland, J. (1991). A computational model of semantic memory impairment: Modality specificity and emergent category specificity. Journal of Experimental Psychology: General, 120(4), 339-357.

Farah, M., & McClelland, J. (1992). Neural network models and cognitive neuropsychology. Psychiatric Annals, 22(3), 148-153.

Feldman, J., & Drasgow, J. (1951). A visual-verbal test for schizophrenia. Psychiatric Quarterly Supplement, 25, 55-64.

Flicker, C., Ferris, S., Crook, T., & Bartus, R. (1987). Implications of memory and language dysfunction in the naming deficit of senile dementia. Brain and Language, 31, 187-200.

Folstein, M., Folstein, S., & McHugh, P. (1975). Mini-mental state: A practical method for grading the cognitive state of patients for the clinician. Journal of Psychiatric Research, 12, 189-198.

Fox, C., & Jones, J. (1998). Uses of Rasch modeling in counseling psychology research. Journal of Counseling Psychology, 45(1), 30-45.

Fuld, P. (1983). Psychometric differentiation of the dementias: An overview. In B. Reisberg (Ed.), Alzheimer's disease: The standard reference (pp. 201-213). New York: The Free Press.

Fuster, J. (2000). The prefrontal cortex of the primate: A synopsis. Psychobiology, 28(2), 125-131.

Galasko, D., Klauber, M., Hofstetter, R., Salmon, D., Lasker, B., & Thal, L. (1990). The Mini-Mental State Examination in the early diagnosis of Alzheimer's disease. Archives of Neurology, 47, 49-52.

Gentner, D. (1989). The mechanisms of analogical reasoning. In S. Vosniadou, & A. Ortony (Eds.), Similarity and analogical reasoning (pp. 199-241). London: Cambridge Univ. Press.

Gewirth, L., Shindler, A., & Hier, D. (1984). Altered patterns of word associations in dementia and aphasia. Brain and Language, 21, 301-317.

Gitomer, D., Curtis, M., Glaser, R., & Lensky, D. (1987). Processing differences as a function of item difficulty in verbal analogy performance. Journal of Educational Psychology, 79, 212-219.

Goel, V., & Dolan, R. (2000). Anatomical segregation of component processes in an inductive inference task. Journal of Cognitive Neuroscience, 12(1), 110-119.

Goldblum, M., Gomez, C., Dalla Barba, G., Boller, F., Deweer, B., Hahn, V., & Dubois, B. (1998). The influence of semantic and perceptual encoding on recognition memory in Alzheimer's disease. Neuropsychologia, 36(8), 717-729.

Goldman, S., Pellegrino, J., Parseghian, P., & Sallis, R. (1982). Developmental and individual differences in verbal analogical reasoning. Child Development, 53(2), 550-559.

Gorham, D. (1956). A proverbs test for clinical and experimental use. Psychological Reports, 1, 1-12.

Goswami, U., & Brown, A. (1989). Melting chocolate and melting snowmen: Analogical reasoning and causal relations. Cognition, 35, 69-95.

Goswami, U. (1991). Analogical reasoning: What develops? A review of research and theory. Child Development, 62, 1-22.

Grady, C., Haxby, J., Horwitz, B., Sundaram, M., Berg, G., Schapiro, M., Friedland, R., & Rapoport, S. (1988). Longitudinal study of the early neuropsychological and cerebral metabolic changes in dementia of the Alzheimer type. Journal of Clinical and Experimental Neuropsychology, 10(5), 576-596.

Graves, R., Carswell, L., & Snow, W.G. (1999). An evaluation of the sensitivity of premorbid IQ estimators for detecting cognitive decline. Psychological Assessment, 11(1), 29-38.

Grossman, M., Jickanin, J., Robinson, K., & D'Esposito, M. (1996). Anomaly judgments of subject-predicate relations in Alzheimer's disease. Brain and Language, 54(2), 216-232.

Hart, R., Kwentus, J., Taylor, J., & Hamer, R. (1988). Productive naming and memory in depression and Alzheimer's type dementia. Archives of Clinical Neuropsychology, 3, 313-322.

Hartman, M. (1991). The use of semantic knowledge in Alzheimer's disease: Evidence for impairments of attention. Neuropsycholgia, 29(3), 213-228.

Hattie, J. (1985). Methodology review: Assessing unidimensionality of tests and items. Applied Psychological Measurement, 9(2). 139-164.

Heidrich, S., & Denney, N. (1994). Does social problem solving differ from other types of problem solving during the adult years? Experimental Aging Research, 20, 105-124.

Heinik, J., & Aharon-Peretz, J. (1993). Abstract thinking in dementia: A preliminary report. Clinical Gerontologist, 12(4), 57-60.

Holyoak, K., & Thagard, P. (1989). Analogical mapping by constraint satisfaction. Cognitive Science, 13, 295-355.

Huff, F., Corkin, S., & Growdon, J. (1986). Semantic impairment and anomia in Alzheimer's disease. Brain and Language, 28, 235-249.

Hughes, J., Graham, N., Patterson, K., & Hodges, J. (1997). Dysgraphia in mild dementia of the Alzheimer's type. Neuropsychologia, 35(4), 533-545.

Jarvik, L. (1988). Aging of the brain: How can we prevent it? The Gerontologist, 28, 739-747.

Johnson-Laird, P., & Byrne, R. (1991). Deduction. Hillsdale, NJ: Erlbaum.

Kaplan, E., Goodglass, H., & Weintraub, S. (1978). The Boston Naming Test. Boston: Kaplan , & Goodglass.

Kaufman, A., & Kaufman, N. (1983). K-ABC. Kaufman Assessment Battery for Children. Circle Pines, MN: American Guidance Service.

Kaufman, A., Kaufman-Packer, J., McLean, J., & Reynolds, C. (1991). Is the pattern of intellectual growth and decline across the adult life span different for men and women? Journal of Clinical Psychology, 47, 801-812.

Keane, M., Ledgeway, T., & Duff, S. (1994). Constraints on analogical mapping: A comparison of three models. Cognitive Science, 18, 387-438.

Koechlin, E., Basso, G., Pietrini, P., Panzer, S., & Grafman, J. (1999). The role of the anterior prefrontal cortex in human cognition. Nature, 399, 148-151.

Koss, E., Haxby, J, DeCarli, C., Schapiro, M., Friedland, R., & Rapoport, S. (1991). Patterns of performance preservation and loss in healthy aging. Developmental Neuropsychology, 7, 99-113.

Lafleche, G., & Albert, M. (1995). Executive function deficits in mild Alzheimer's disease. Neuropsychology, 9(3), 313-320.

Larrabee, G., Largen, J., & Levin, H. (1985). Sensitivity of age-decline resistant ("Hold") WAIS subtests to Alzheimer's disease. Journal of Clinical and Experimental Neuropsychology, 7, 497-504.

Lezak, M. (1995). Neuropsychological assessment, 3rd edition. New York: Oxford University Press.

Martin, A., & Fedio, P. (1983). Word production and comprehension in Alzheimer's disease: The breakdown of semantic knowledge. Brain and Language, 19, 121-141.

Masterson, J., Evans, L., & Aloia, M. (1993). Verbal analogical reasoning in children with language-learning disabilities. Journal of Speech and Hearing Research, 36, 76-82.

McCarthy, R., & Warrington, E. (1988). Evidence for modality-specific meaning systems in the brain. Nature, 334, 428-430.

McKhann, G., Drachman, D., Folstein, M., Katzman, R., Price, D., & Stadlan, E. (1984). Clinical diagnosis of Alzheimer's disease: Report of the NINCDS-ADRDA work group under the auspices of Department of Health and Human Services Task Force on Alzheimer's disease. Neurology, 34, 939-944.

Measso, G., Zappala, G., Cavarzeran, R., Crook, T., Romani, L., Pirozzolo, F., Grigoletto, F., Amaducci, L., Massari, D., & Lebowitz, B. (1993). Raven's coloured progressive matrices: A normative study of a random sample of healthy adults. Acta Neurologica Scandinavica, 88, 70-74.

Morrison, R., Krawczyk, D., Knowlton, B., Holyoak, K., Boone, K., Chow, T., & Mishkin, F. (2000, March). Relational reasoning and semantic inhibition in human prefrontal cortex. Poster presented at the 10th annual Frontal Lobes conference of the Rotman Research Institute, Toronto, Ontario.

Nebes, R. (1989). Semantic memory in Alzheimer's disease. Psychological Bulletin, 106(3), 377-394.

Nebes, R., Brady, C., & Huff, F. (1989). Automatic and attentional mechanisms of semantic priming in Alzheimer's disease. Journal of Clinical and Experimental Neuropsychology, 11, 219-230.

Nebes, R., Martin, D., & Horn, L. (1984). Sparing of semantic memory in Alzheimer's disease. Journal of Abnormal Psychology, 93, 321-330.

Niederehe, G. (1986). Depression and memory impairment in the aged. In L. Poon , & T. Crook (Eds.), Handbook for clinical memory assessment of older adults (pp. 226-237). Washington, DC: American Psychological Association.

Nippold, M., Erskine, B., & Freed, D. (1988). Proportional and functional analogical reasoning in normal and language-impaired children. Journal of Speech and Hearning Disorders, 53, 440-448.

Ober, B., Dronkers, N., Koss, E., Delis, D., & Friedland, R. (1986). Retrieval from semantic memory in Alzheimer-type dementia. Journal of Clinical and Experimental Neuropsychology, 8, 75-92.

Ober, B., Shenaut, G., Jagust, W., & Stillman, R. (1991). Automatic semantic priming with varying types of category relationships in Alzheimer's disease and normal aging. Psychology and Aging, 6, 647-660.

Ober, B., Shenaut, G., & Reed, B. (1995). Assessment of associative relations in Alzheimer's disease: Evidence for preservation of semantic memory. Aging and Cognition, 2(4), 254-267.

Orme, J. (1966). Hypothetically true norms for the progressive matrices test. Human Development, 9, 222-230.

Paivio, A. (1991). Dual coding theory: Retrospect and current status. Canadian Journal of Psychology, 45, 255-287.

Panek, P., & Stoner, S. (1980). Age differences on Raven's Coloured Progressive Matrices. Perceptual and Motor Skills, 50, 977-978.

Potter, M., & Faulconer, B. (1975). Time to understand words and pictures. Nature, 253, 437-438.

Prabhakaran, V., Narayanan, K., Zhao, Z., & Gabrieli, J. (2000). Integration of diverse information in working memory within the frontal lobe. Nature Neuroscience, 3(1), 85-90.

Raskin, A. (1986). Partialing out the effects of depression and age on cognitive functions: Experimental data and methodologic issues. In L. Poon , & T. Crook (Eds.), Handbook for clinical memory assessment of older adults (pp. 244-255). Washington, DC: American Psychological Association.

Raven, J.C., Court, J., & Raven, J. (1976). Manual for Raven's Progressive Matrices. London: H.K. Lewis , & Co. Ltd.

Raven, J.C., Court, J., & Raven, J. (1990). Manual for Raven's Progressive Matrices and Vocabulary Scales. Oxford, England: Oxford Psychologist's press.

Raven, J.C. (no date). Raven's Progressive Matrices. Examination kit. Los Angeles: Western Psychological Services.

Reisberg, B., Ferris, S., Borenstein, J., Sinaiko, E., de Leon, M., & Buttinger, C. (1986). Assessment of presenting symptoms. In L. Poon , & T. Crook (Eds.), Handbook

for clinical memory assessment of older adults (pp. 108-128). Washington, DC: American Psychological Association.

Riddoch, M., Humphreys, G., Coltheart, M., & Funnell, E. (1988). Semantic systems or system? Neuropsychological evidence re-examined. Cognitive Neuropsychology, 5, 3-25.

Rosen, W. (1983). Clinical and neuropsychological assessment of Alzheimer's disease. In R. Mayeux, & W. Rosen (Eds.), The dementias (pp. 51-64). New York: Raven Press.

Ross, T., Lichtenberg, P., & Christensen, B. (1995). Normative data on the Boston Naming Test for elderly adults in a demographicaly diverse medical sample. The Clinical Neuropsychologist, 9(4), 321-325.

Ryan, J., & Paolo, A. (1992). A screening procedure for estimating premorbid intelligence in the elderly. The Clinical Neuropsychologist, 6, 53-62.

Sabe, L., Jason, L., Juejati, M., Leiguarda, R., & Starkstein, S. (1993). Sensitivity and specificity of the Mini-Mental State Exam in the diagnosis of dementia. Behavioural Neurology, 6, 207-210.

Salmon, D., & Butters, N. (1992). Neuropsychologic assessment of dementia in the elderly. In R. Katzman , & J. W. Rowe (Eds.), Principles of geriatric neurology (pp. 144-163). Philadelphia: F.A. Davis Company.

Salthouse, T. (2000). Item analyses of age relations on reasoning tests. Psychology and Aging, 15(1), 3-8.

Salthouse, T. (1992). Reasoning and spatial abilities. In F. Craik , & T. Salthouse (Eds.), The handbook of aging and cognition (pp. 167-211). Hillsdale, New Jersey: Lawrence Erlbaum Associates, Publishers.

Salthouse, T., & Skovronek, E. (1992). Within-context assessment of age differences in working memory. Journals of Gerontology, 47(3), P110-P120.

Sawrie, S., Chelune, G., Naugle, R., & Luders, H. (1996). Empirical methods for assessing meaningful neuropsychological change following epilepsy surgery. Journal of the International Neuropsychological Society, 2(6), 556-564.

Schmitt, N. (1996). Uses and abuses of coefficient alpha. Psychologica, 8, 350-353.

Schultz, N., Kaye, D., & Hoyer, W. (1980). Intelligence and spontaneous flexibility in adulthood and old age. Intelligence, 4, 219-231.

Schwartz, M., Marin, O., & Saffran, E. (1979). Dissociation of language function in dementia: A case study. Brain and Language, 7, 277-306.

Shallice, T. (1993). Selective impairments of knowledge. In T. Shallice (Ed.), From neuropsychology to mental structure (pp. 269-306). New York: Cambridge University Press.

Shuttleworth, E., & Huber, S. (1989). The picture absurdities test in the evaluation of dementia. Brain and Cognition, 11, 50-59.

Spellman, B., & Holyoak, K. (1996). Pragmatics in analogical mapping. Cognitive Psychology, 31, 307-346.

Stebbins, G., Wilson, R., Gilley, D., Bernard, B., & Fox, J. (1990). Use of the National Adult Reading Test to estimate premorbid IQ in dementia. The Clinical Neuropsychologist, 4, 18-24.

Sternberg, R. (1977). Component processes in analogical reasoning. Psychological Review, 84, 353-378.

Teng, E., & Chui, H. (1987). The modified Mini-Mental State (3MS) Examination. Journal of Clinical Psychiatry, 48, 314-318.

Terman, L., & Merrill, M. (1973). Standford-Binet Intelligence Scale. Manual for the Third Revision, Form L-M. Boston: Houghton-Mifflin Co.

Tombaugh, T., McDowell, I., Kristjansson, B., & Hubley, A. (1996). Mini-Mental State Examination (MMSE) and the Modified MMSE (3MS): A psychometric comparison and normative data. Psychological Assessment, 8, 48-59.

Tombaugh, T., & McIntyre, N. (1992). The Mini-Mental State Examination: A comprehensive review. Journal of the American Geriatrics Society, 40, 922-935.

Tow, P. (1955). Personality changes following frontal leucotomy. London: Oxford University Press.

Tuokko, H. (1993). Psychosocial evaluation and management of the Alzheimer's patient. In R.W. Parks, R.F. Zec , & R.S. Wilson (Eds.), Neuropsychology of Alzheimer's disease and other dementias (pp. 565-588). New York: Oxford University Press.

Tuokko, H., & Hadjistavropoulos, T. (1998). An assessment guide to geriatric neuropsychology. London: Lawrence Erlbaum Associates.

Villardita, C. (1985). Raven's Colored Progressive Matrices and intellectual impairment in patients with focal brain damage. Cortex, 21, 627-634.

Waltz, J., Knowlton, B., Holyoak, K., Boone, K., Mishkin, F., de-Menezes-Santos, M., Thomas, C., & Miller, B. (1999). A system for relational reasoning in human prefrontal cortex. Psychological-Science, 10(2), 119-125.

Warrington, E. (1975). The selective impairment of semantic memory. Quarterly Journal of Experimental Psychology, 27, 635-657.

Wechsler, D. (1997). Wechsler Adult Intelligence Scale – III. Toronto: The Psychological Corporation: Harcourt Brace , & Company.

Wechsler, D. (1981). Wechsler Adult Intelligence Scale-Revised. New York: The Psychological Corporation: Harcourt Brace Jovanovich, Inc.

Weingartner, H., Kawas, C., Rawlings, R., & Shapiro, M. (1993). Changes in semantic memory in early stage Alzheimer's disease patients. Gerontologist, 33(5), 637-643.

Welsh, K., Butters, N., Hughes, J., Mohs, R., & Heyman, A. (1991). Detection of abnormal memory decline in mild cases of Alzheimer's disease using CERAD neuropsychological measures. Archives of Neurology, 48, 278-281.

Yesavage, J., Brink, T., Rose, T., Lum, O., Huang, V., Adey, M., & Leirer, V. (1983). Development and validation of a geriatric depression rating scale: A preliminary report. Journal of Psychiatric Research, 17, 37-49.

Yeudall, L., Fromm, D., Reddon, J., & Stefanyk, W. (1986). Normative data stratified by age and sex for 12 neuropsychological tests. Journal of Clinical Psychology, 42, 918-946.

Zec, R. (1993). Neuropsychological functioning in Alzheimer's disease. In R. Parks, R. Zec , & R. Wilson (Eds.), Neuropsychology of Alzheimer's disease and other dementias (pp. 3-80). New York: Oxford University Press.

Table 1

Demographic Data for the AD Group and the Nondemented Controls

	AD Group N= 47	**Nondemented Groups** n=47	n=60
Age (years)			
Mean	78.4	77.7	76.65
Sd	6.9	6.6	7.31
Range	57 - 91	58 - 89	58 - 94
Education (years)			
Mean	11.85	12.68	13.17
Sd	2.92	2.63	2.77
Range	6 - 17	7 - 18	7 - 18
Gender			
Female	70%	70%	68%
Male	30%	30%	32%

Table 2

Mean Premorbid VIQ Estimates for the Matched AD and Nondemented Groups

	AD	Nondemented
NART		
Mean	106.3	111.34
sd	12.35	11.52
Range	75-127	84-129
Barona Index		
Mean	105.23	107.94
sd	9.26	8.26
Range	87 - 121	84 - 121

Table 3

Screening Measure Performance of AD (n = 47) and Nondemented (n = 47) Groups

	Alzheimer's		Nondemented	
	Mean	sd	Mean	sd
Geriatric Depression Scale	7.15	(3.77)	4.55	(2.9)
Boston Naming Test	24th	(29.06)	58th	(25.16)
MMSE	16th	(14.94)	57th	(25.33)
3MS	25th	(24.49)	70th	(22.9)

Table 4

GAT Items Eliminated, Eigenvalues and Explained Variance for the First Two Factors for GAT Problem Subgroups

	Eliminated Items	Eigenvalues		Explained Variance	
		Factor 1	Factor 2	Factor 1	Factor 2
GAT (Version 1)					
Picture Problems (Set A)	# 2, 6, 14, 16, 22	4.48	1.95	23.59%	10.27%
Word Problems (Set B)	# 5, 8, 14, 16, 23	5.86	2.08	30.82%	10.96%
GAT (Version 2)					
Picture Problems (Set B)	# 5, 8, 14, 16, 23	5.64	1.66	29.72%	8.77%
Word Problems (Set A)	# 2, 6, 14, 16, 22	4.55	1.88	23.96%	9.93%

Table 5

Means, Standard Deviations, and Range of Scores on GAT Sections and Versions for the AD and Nondemented (ND) Groups

	Mean		SD		Range	
	AD	ND	AD	ND	AD	ND
GAT Version 1 (38 items)	20.56	31.41	6.5	4.5	9 - 32	19 - 37
Word Problems (19 items)	10.0	16.05	3.8	2.6	4 - 17	10 - 19
Picture Problems (19 items)	10.56	15.36	3.2	2.7	5 - 16	9 - 19
GAT Version 2 (38 items)	23.18	29.8	6.9	6.6	1 - 34	18 - 38
Word Problems (19 items)	13.86	16.2	3.8	2.7	1 - 19	10 - 19
Picture Problems (19 items)	9.32	13.6	3.8	4.4	0 - 17	5 - 19

Table 6

Convergent and Divergent Validity of the GAT

	GAT (raw score) 38 items			Matrix Reasoning (age scaled)			RCPM (%ile)			Similarities (age scaled)			GDS (raw score)		
	Both Groups	**AD**	**CTL**	**Both Groups**	**AD**	**CTL**	**Both Groups**	**AD**	**CTL**	**Both Groups**	**AD**	**CTL**	**Both Groups**	**AD**	**CTL**
GAT (V1)	1.0	1.0	1.0	.63**	.59**	.53*	.68**	.79**	.35*	.76**	.67**	.42*	-.11	.15	-.01
MR	.63**	.59**	.53*	1.0	1.0	1.0	.64**	.64*	.67**	.55**	.47*	.42*	-.17	-.04	-.11
RCPM	.68**	.79**	.35*	.64**	.64*	.67**	1.0	1.0	1.0	.55**	.49*	.07	-.28	-.12	-.27
Similarities	.76**	.67**	.42*	.55**	.47*	.42*	.55**	.49*	.07	1.0	1.0	1.0	.01	.27	.24
GDS	-.11	.15	-.01	-.17	-.04	-.11	-.28	-.12	-.27	.01	.27	.24	1.0	1.0	1.0
BNT	.75**	.63**	.38	.50**	.44*	.32	.64**	.45*	-.001	.72**	.69**	.35	-.03	.31	.12
MMSE	.71**	.56**	.30	.62**	.57**	.55**	.66**	.58**	.35	.68**	.39	.47*	-.15	.22	.03
3MS	.75**	.59**	.35	.60**	.62**	.46*	.66**	.65**	.09	.72**	.52**	.50*	-.08	.19	.26
GAT (V2)	1.0	1.0	1.0	.70**	.59**	.65**	.69**	.60**	.75**	.71**	.49*	.78**	-.37	-.04	-.43
MR	.70**	.59**	.65**	1.0	1.0	1.0	.74**	.67**	.75**	.72**	.53*	.80**	-.50	-.35	-.41
RCPM	.69**	.60**	.75**	.74**	.67**	.75**	1.0	1.0	1.0	.66**	.61**	.65**	-.41	-.28	-.32
Similarities	.71**	.49*	.78**	.72**	.53*	.80**	.66**	.61**	.65**	1.0	1.0	1.0	-.58	-.48	-.49
GDS	-.37	-.04	-.43	-.50	-.35	-.41	-.41	-.28	-.32	-.58	-.48	-.49	1.0	1.0	1.0
BNT	.55**	.21	.72**	.37**	-.04	.52**	.53**	.12	.59**	.50**	.26	.54**	- .42**	-.19	-.44
MMSE	.54**	.10	.53**	.49**	.16	.35	.56**	.24	.39	.51**	.31	.38	- .38**	.11	-.34
3MS	.59**	.18	.70**	.52**	.16	.47*	.49**	-.07	.63**	.52**	.22	.49*	- .32*	.18	-.40

** Correlation significant at .01 level (2-tailed)

* Correlation significant at .05 level (2-tailed)

Table 7

Sensitivity and Specificity of the GAT at Various Cut-Scores

	GAT (Version 1)		GAT (Version 2)	
Cut-Scores	**Sensitivity**	**Specificity**	**Sensitivity**	**Specificity**
24	72%	95%	55%	72%
25	76%	95%	64%	72%
26	76%	91%	64%	72%
27	80%	86%	73%	72%
28	80%	77%	82%	68%
29	88%	68%	86%	64%
30	88%	50%	91%	60%

Table 8

Sensitivity and Specificity of Abstract Thinking and Reasoning Measures

Test	Optimal Cut Score	Sensitivity	Specificity
GAT (Version 1)	27 (raw)	80%	86%
GAT (Version 2)	28 (raw)	82%	68%
RCPM	29 (raw)	83%	62%
Matrix Reasoning	11 (age scaled)	72%	62%
Similarities	11 (age scaled)	74%	62%

Table 9

Positive Predictive Values of Reasoning Measures

Test	Positive Predictive Value
GAT (Version 1)	86%
GAT (Version 2)	69%
RCPM	68%
Matrix Reasoning	66%
Similarities	65%

Table 10

Post-test Probability Calculations for the GAT

	Calculation	GAT (Version 1)	GAT (Version 2)
Likelihood Ratio (LR)	$\frac{\text{Sensitivity}}{1 - \text{specificity}}$	5.7	2.56
Prevalence	$\frac{a + c}{a + b + c + d}$	25/47 = .53	22/47 = .47
Pre-test Odds	$\frac{\text{prevalence}}{1 - \text{prevalence}}$	1.1	.89
Post-test Odds	pretest odds x LR	6.3	2.3
Post-test Probability	$\frac{\text{post-test odds}}{\text{post-test odds} + 1}$	86%	70%

Note: a = the number of AD patients scoring below the cut-off value; b = the number of nondemented scoring below the cut-off value; c = the number of AD patients scoring above the cut-off value; and, d = the number of nondemented scoring above the cut-off value.

Table 11

Likelihood Ratios and Pre- and Post-test Probabilities of Detecting AD

Test	Likelihood Ratio	Pre-test Probability	Post-test Probability
RCPM	2.18	50%	69%
Matrix Reasoning	1.89	50%	65%
Similarities	1.94	50%	66%

Table 12

WAIS-III Matrix Reasoning and RCPM Descriptive Statistics for the AD and Nondemented Groups

	AD Participants (n = 47)	Nondemented (n = 47)
WAIS-III Matrix Reasoning		
Mean (age-scaled score)	9.4	12.47
Standard Deviation	3.06	3.28
Range	4 - 16	6 - 19
RCPM		
Mean (percentile score)	64th	87th
Standard Deviation	27.25	11.94
Range	8th - 95th	55th - 95th

Table 13

Percentages of AD and Nondemented Participants Making Different Error Types Most Commonly on GAT Word and Picture Problems

Error Type	AD		Nondemented	
	Picture	Words	Picture	Words
Associative	55.3%	61.7%	63.8%	46.8%
Category	10.6%	14.9%	2.1%	4.3%
Appearance Match	14.9%	4.3%	10.6%	10.6%
Miscellaneous	0%	0%	0%	2.1%
Tie	19.2%	19.2%	23.4%	36.2%

Table 14

Percentages of AD and Nondemented Making Different RCPM Error Types

RCPM Error Type	AD	Nondemented
B-match	25.5%	21%
C-match	51%	34%
Ties	23.5%	45%

Figure 1. Parallel distributed processing model of semantic memory of Farah and McClelland.

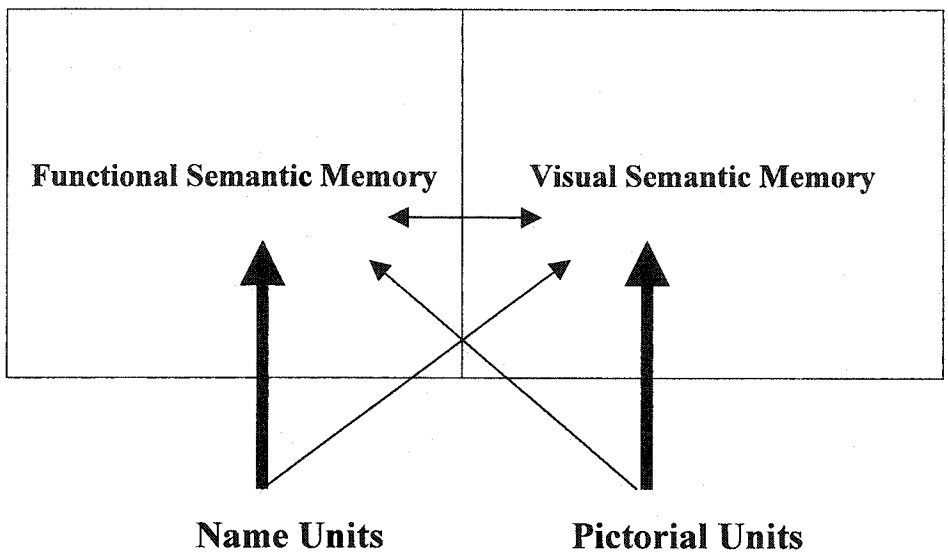


Figure 2. Range and distribution of Boston Naming Test percentile scores as a function of group membership.

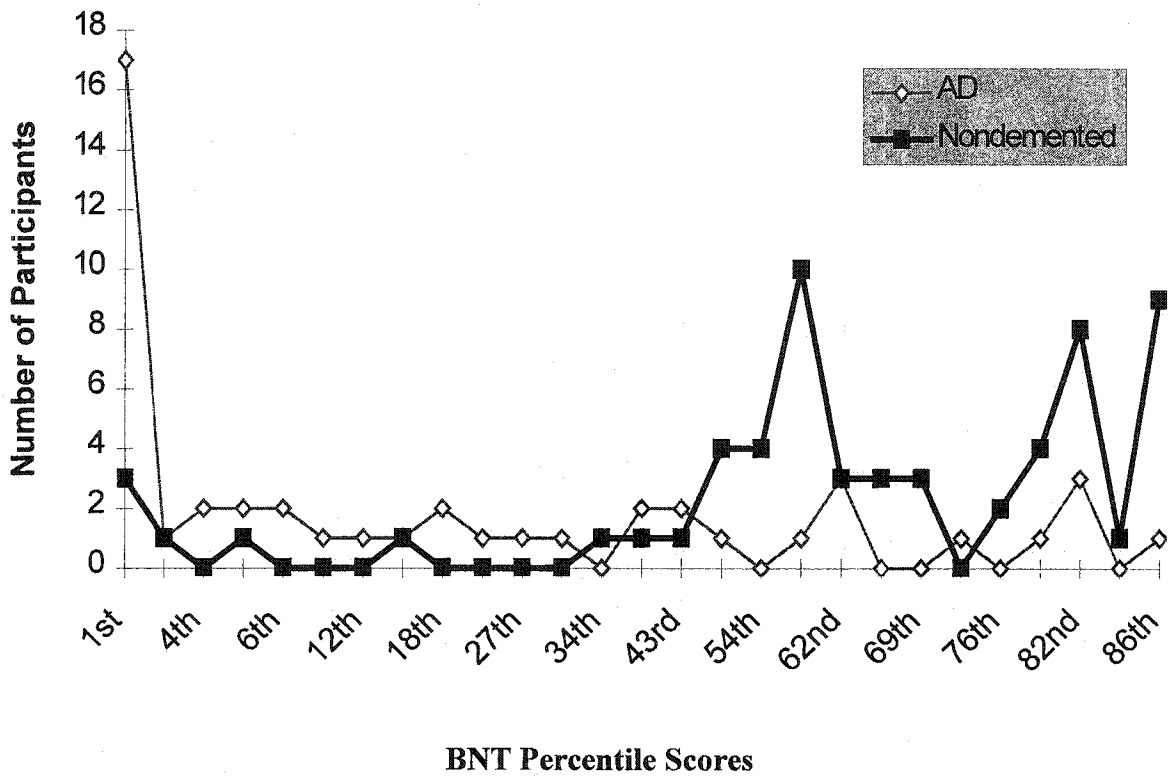


Figure 3. Boxplot diagrams of the GAT scores for the AD and nondemented participants.

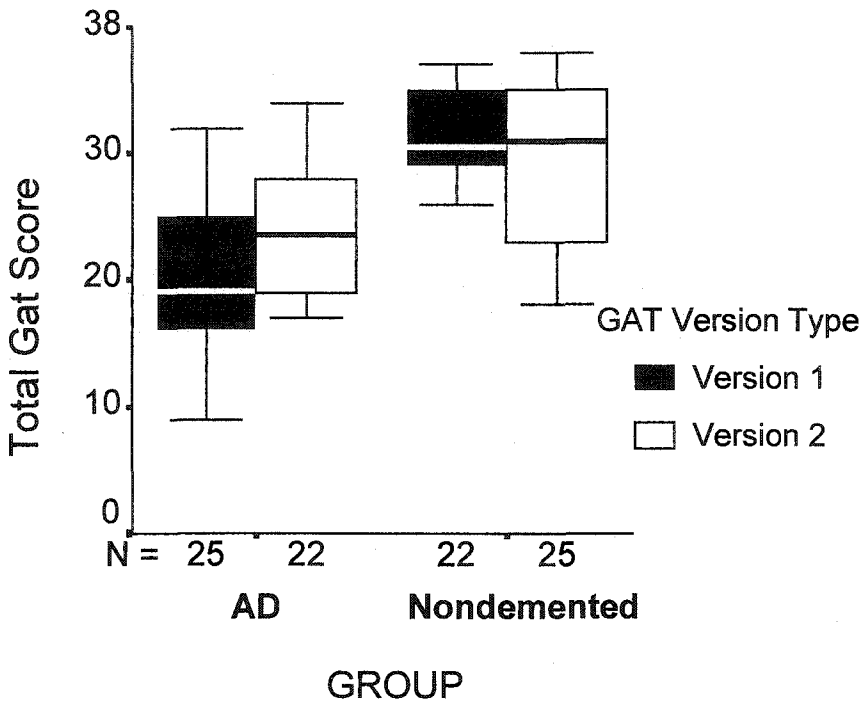
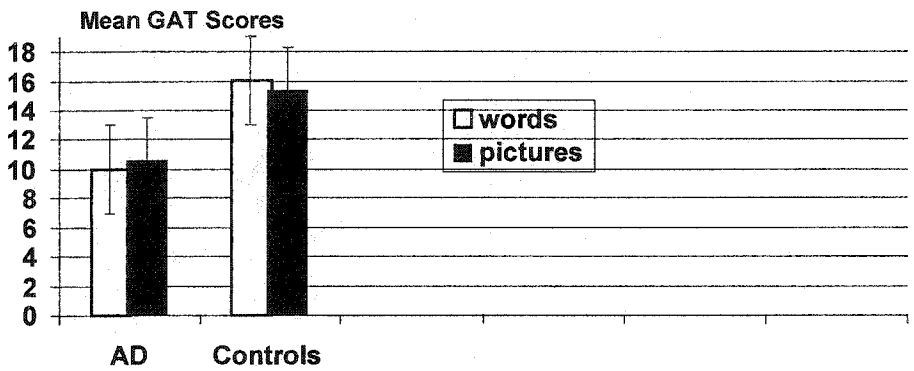
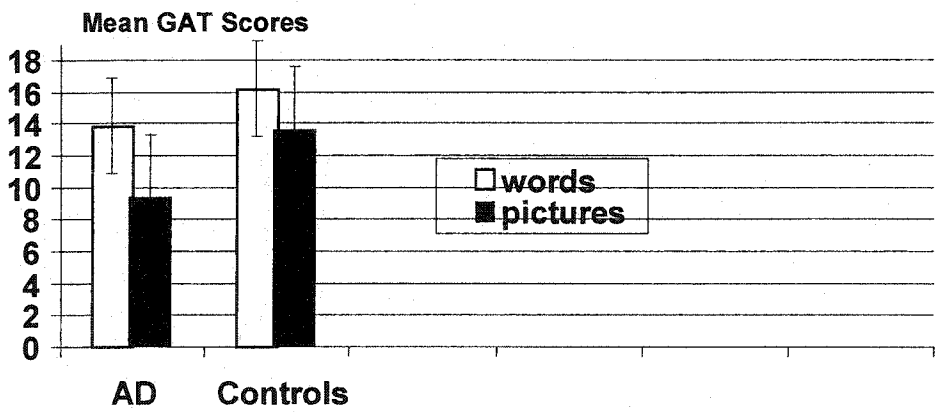


Figure 4. Mean GAT 19-item picture and 19-item word problem scores as a function of group membership and GAT version type

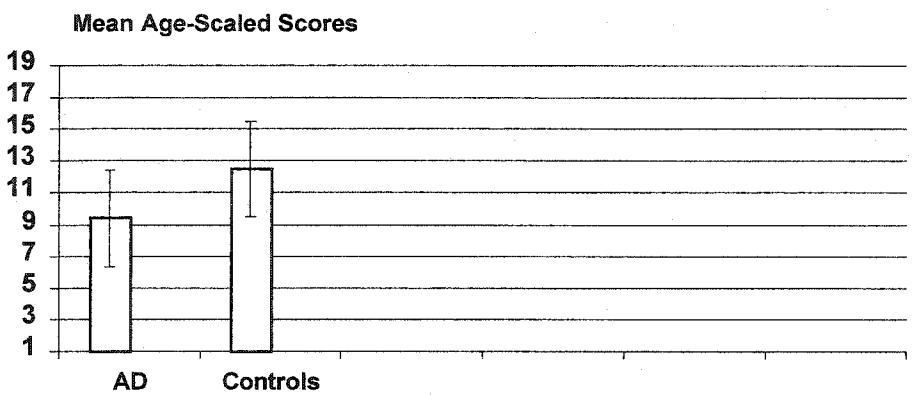


GAT Version 1



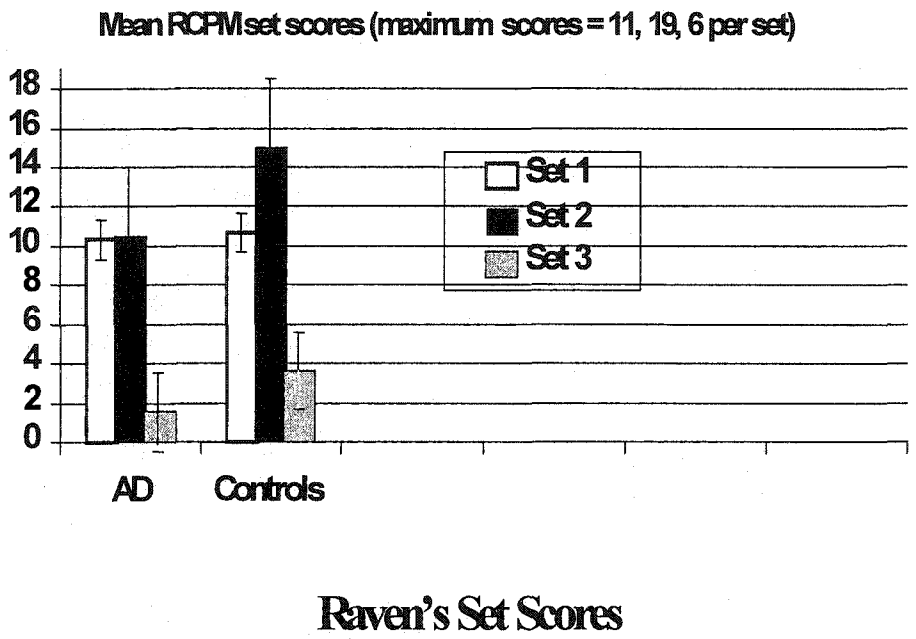
GAT Version 2

Figure 5. Mean WAIS-III Similarities age-scaled scores as a function of group.



WAIS-III Similarities

Figure 6. Mean RCPM set scores for nondemented and AD groups.

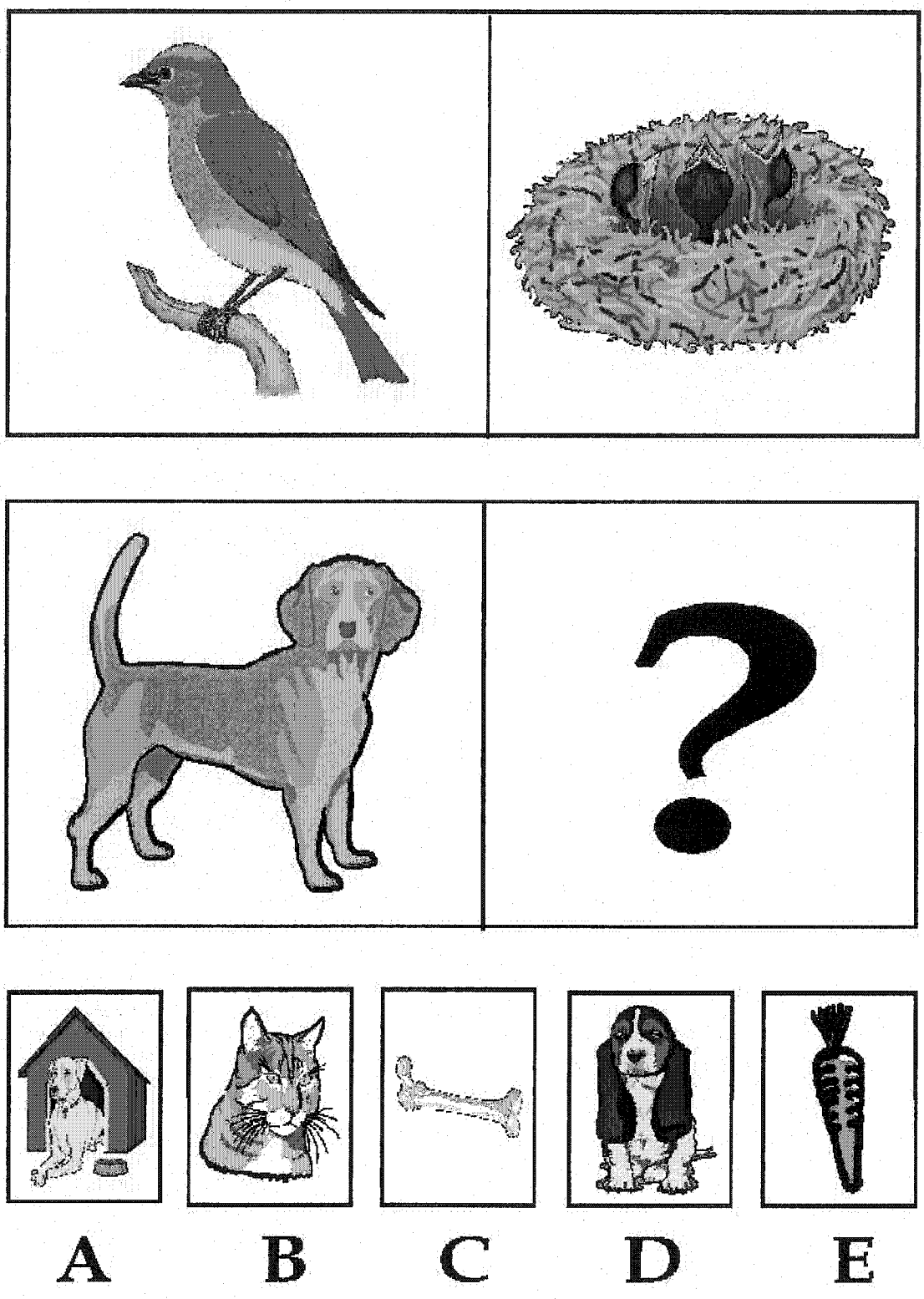


APPENDIX A

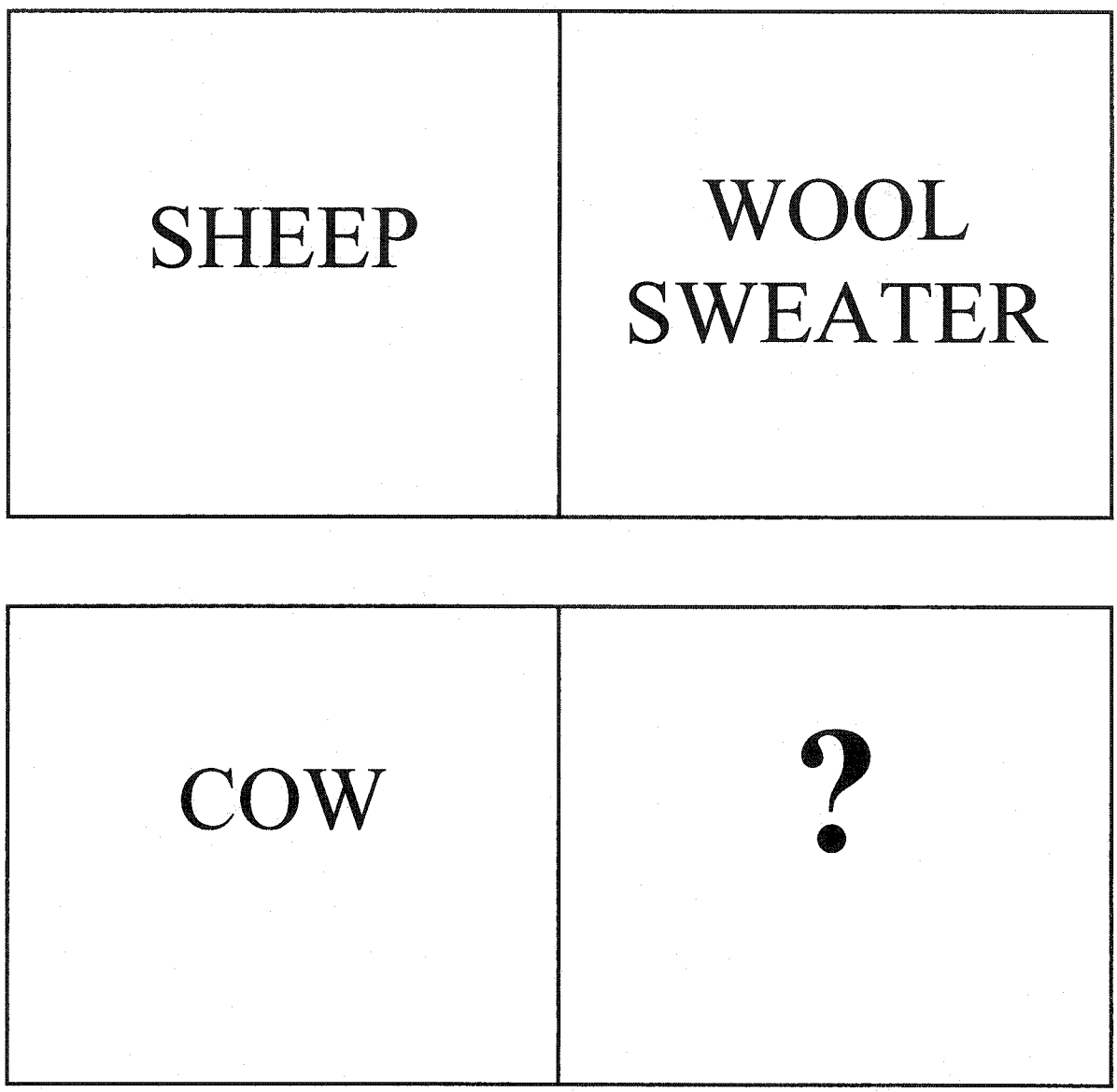
GORANSON ANALOGY TEST

(GAT: Version 1)

Test Developer:	Tamara Goranson, M.A. Ph.D. Candidate in Clinical Neuropsychology Department of Psychology University of Victoria Phone: (519) 721-7525 e-mail: tgoranso@uvic.ca
Ph.D. Supervisor:	Roger Graves, Ph.D., R. Psych. Department of Psychology University of Victoria Phone: (250) 721-7539 Fax: (250) 721-8929 e-mail: rgraves@uvic.ca
University Affiliation:	Department of Psychology University of Victoria P.O. Box 3050 Victoria, B.C., CANADA V8W 3P5

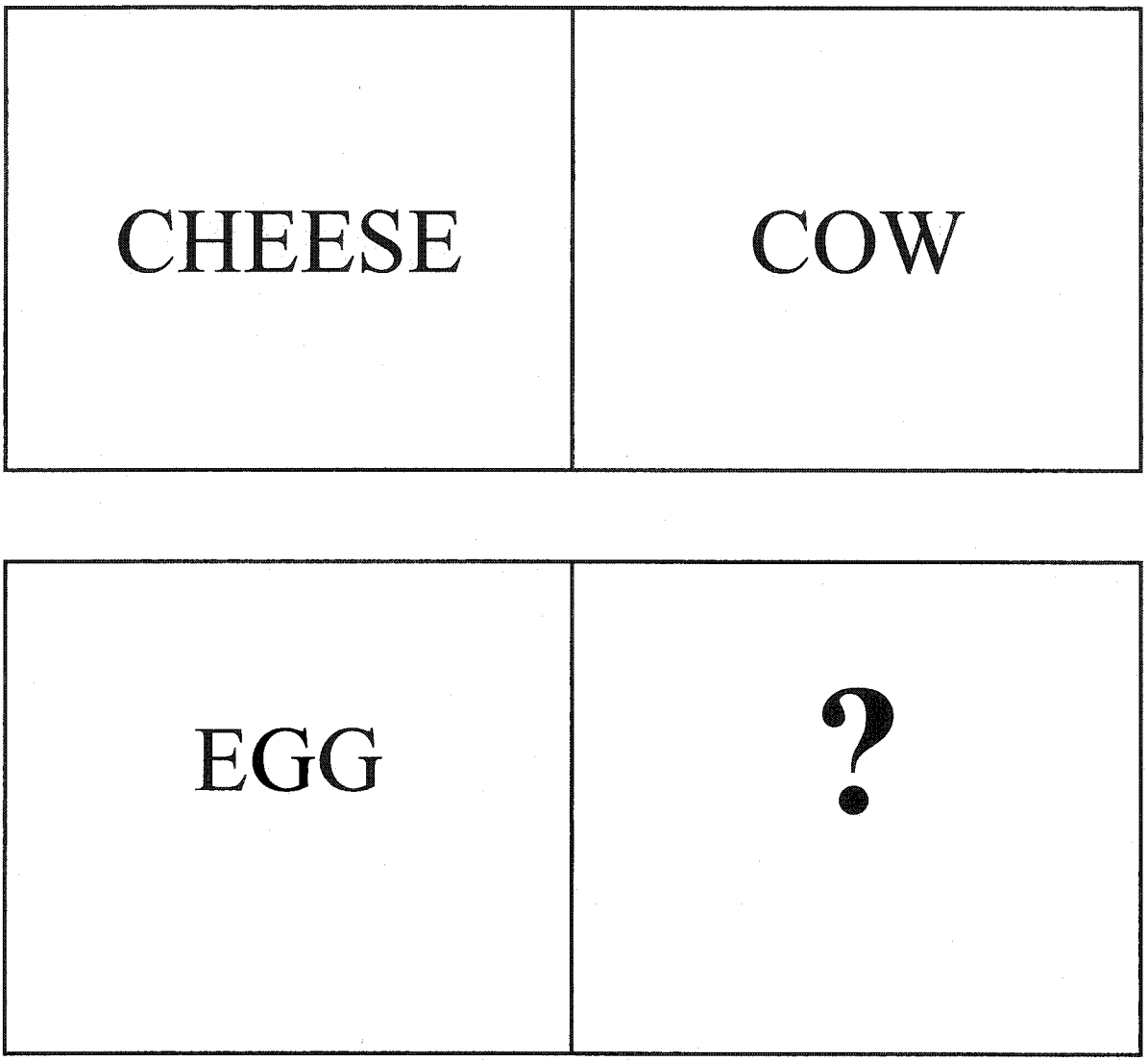


Practice Problem

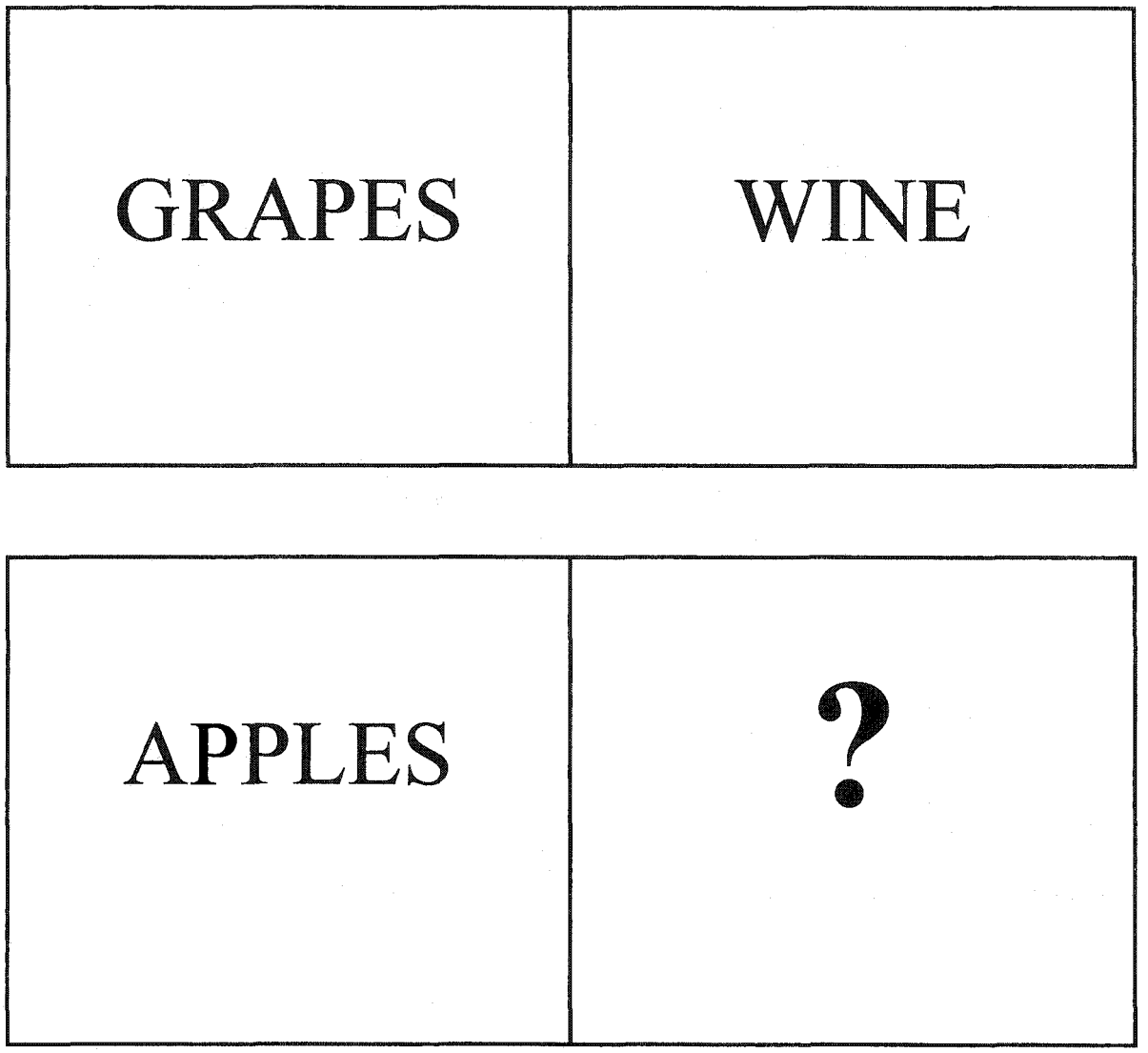


A	B	C	D	E
HORSE	CALF	GARBAGE CAN	LEATHER BOOT	MILK

PRACTICE ITEM

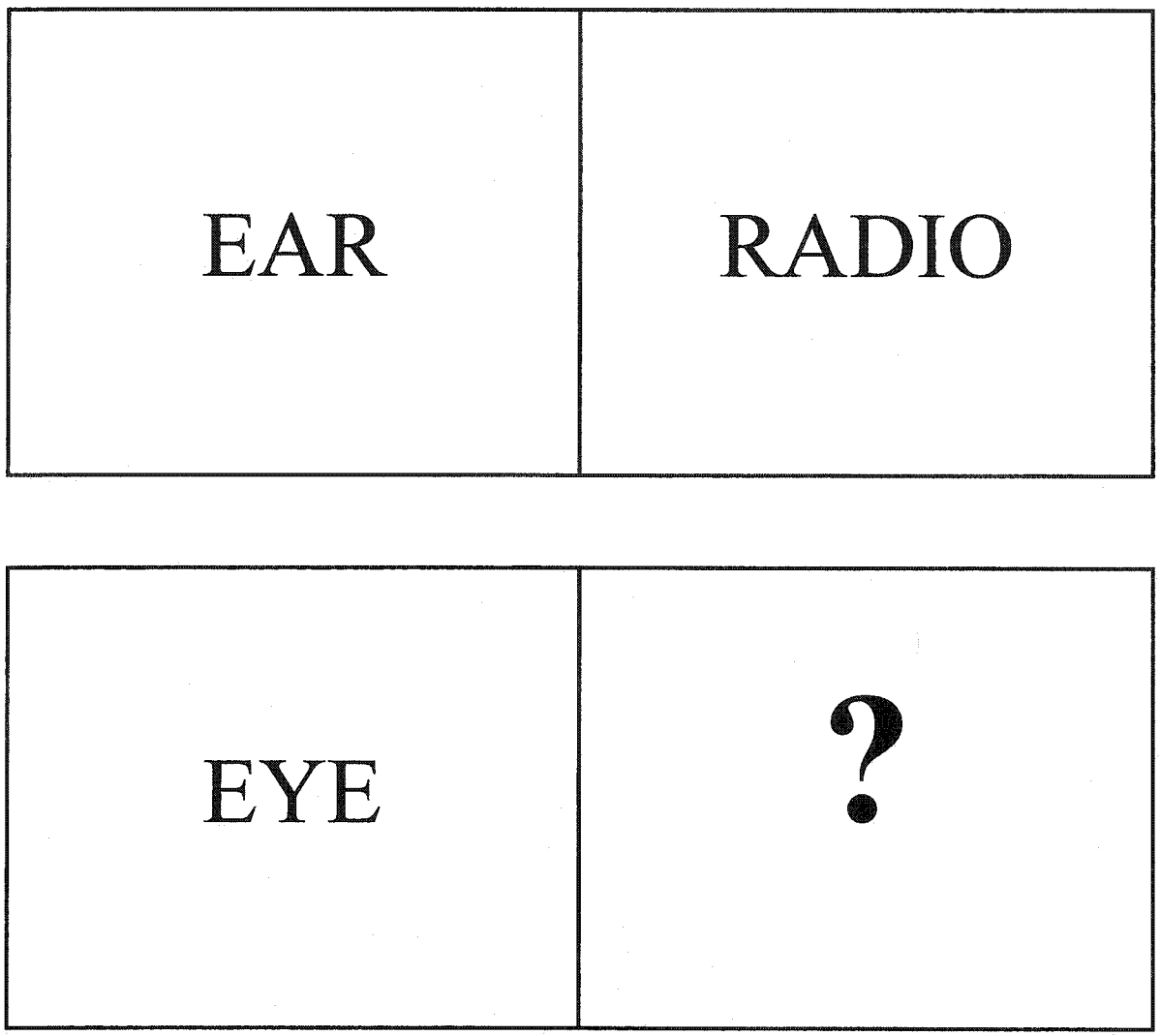


GAT (Version 1) Set B Question 1

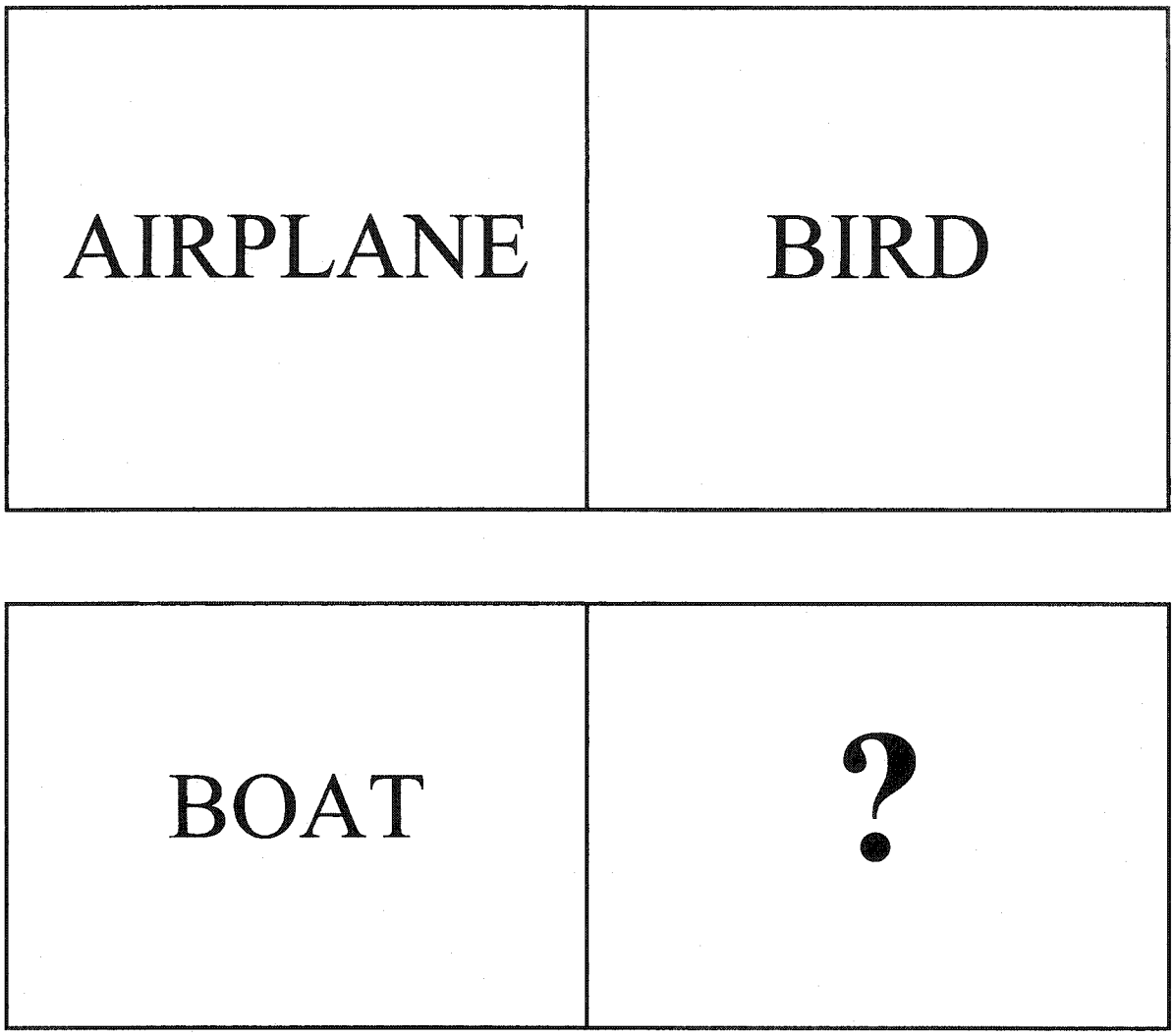


BANANA	TOMATO	BOX	PIE	ORANGES
A	B	C	D	E

GAT (Version 1) Set B Question 2

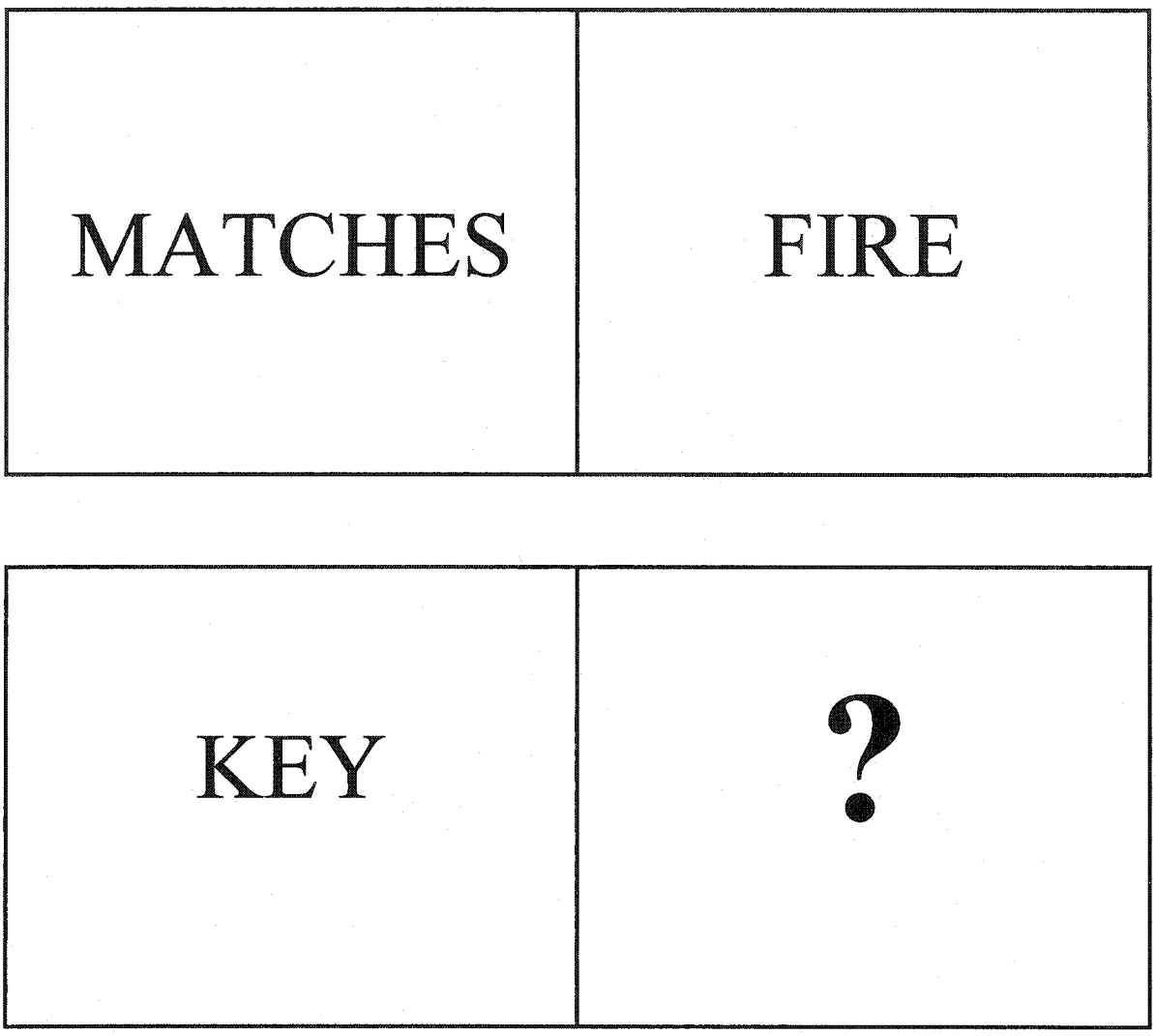


GAT (Version 1) Set B Question 3



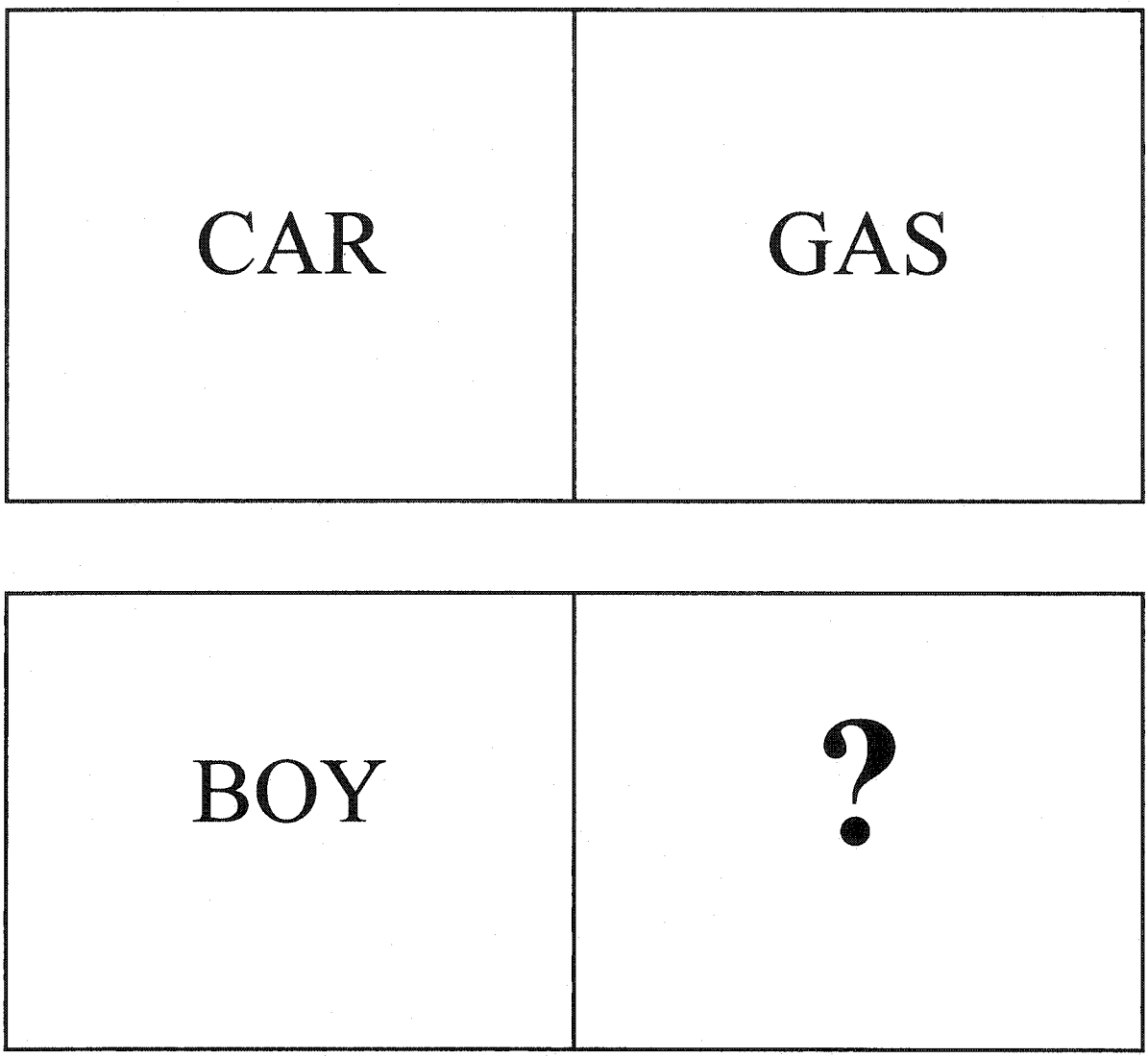
SAILBOAT	BALL	FISH	WATER	TRAIN
A	B	C	D	E

GAT (Version 1) Set B Question 4



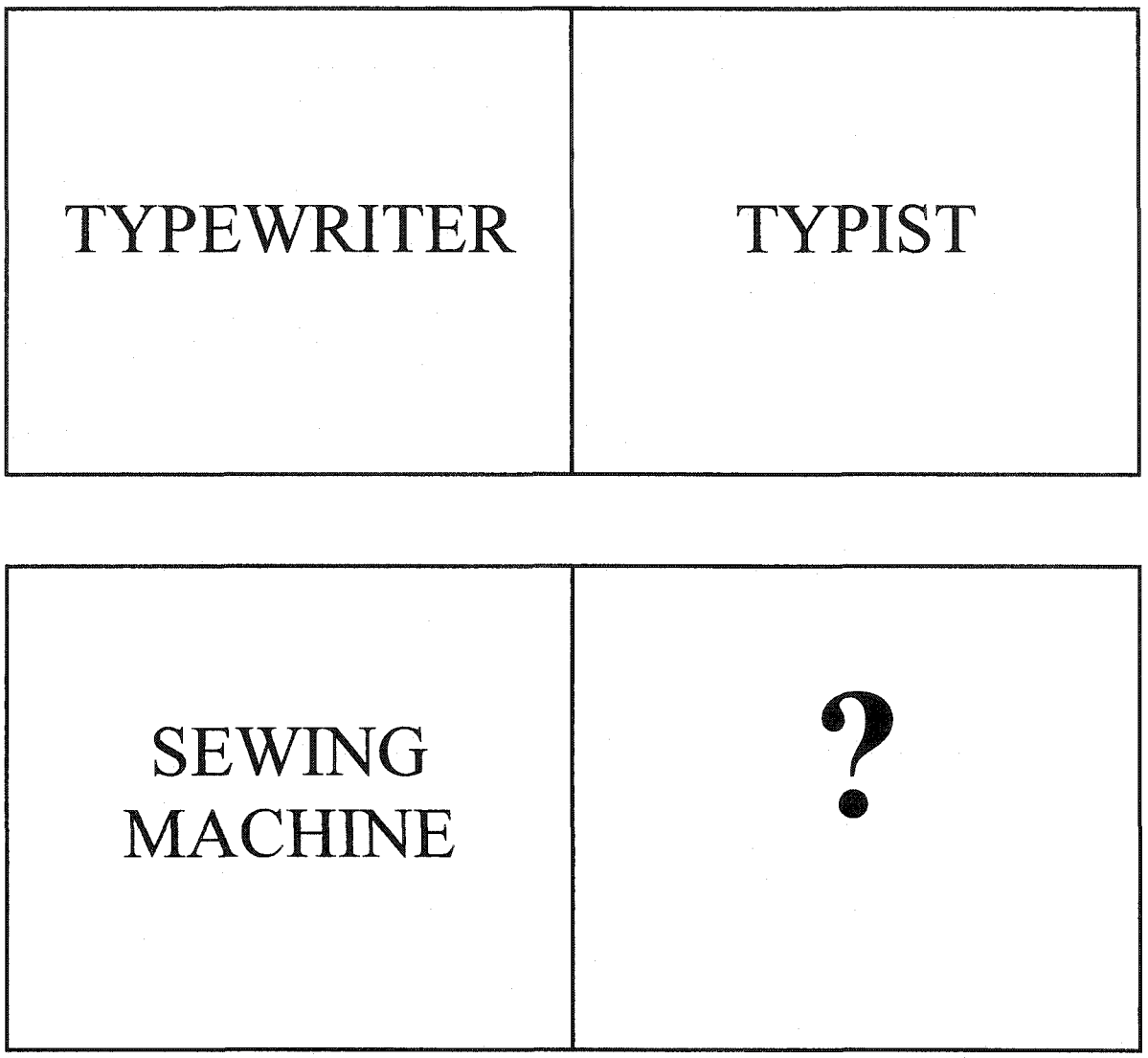
A	B	C	D	E
DOOR	KNIFE	PEAR	CAR	LOCK

GAT (Version 1) Set B Question 5

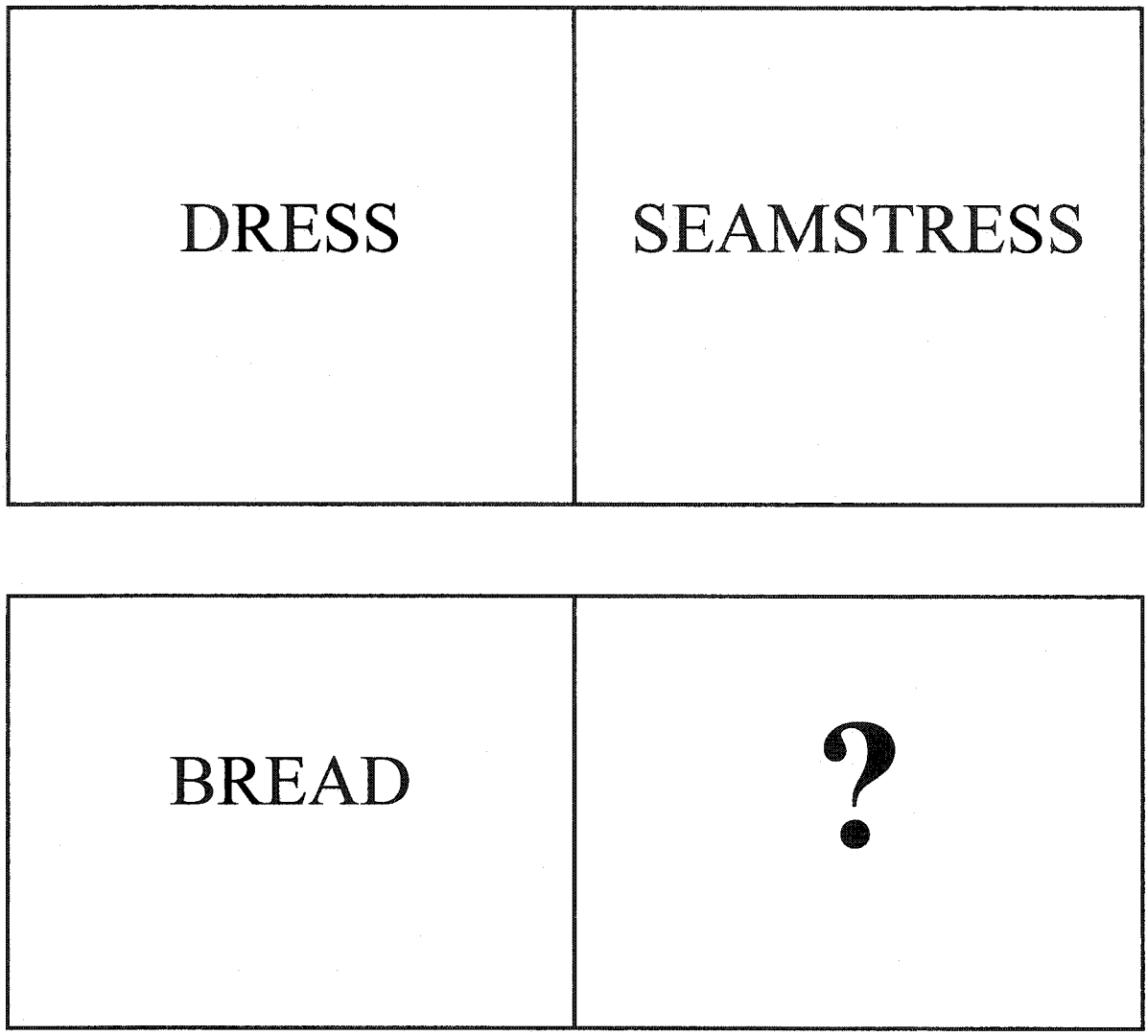


A	B	C	D	E
MAN	TREE	FOOD	GIRL	CHILD

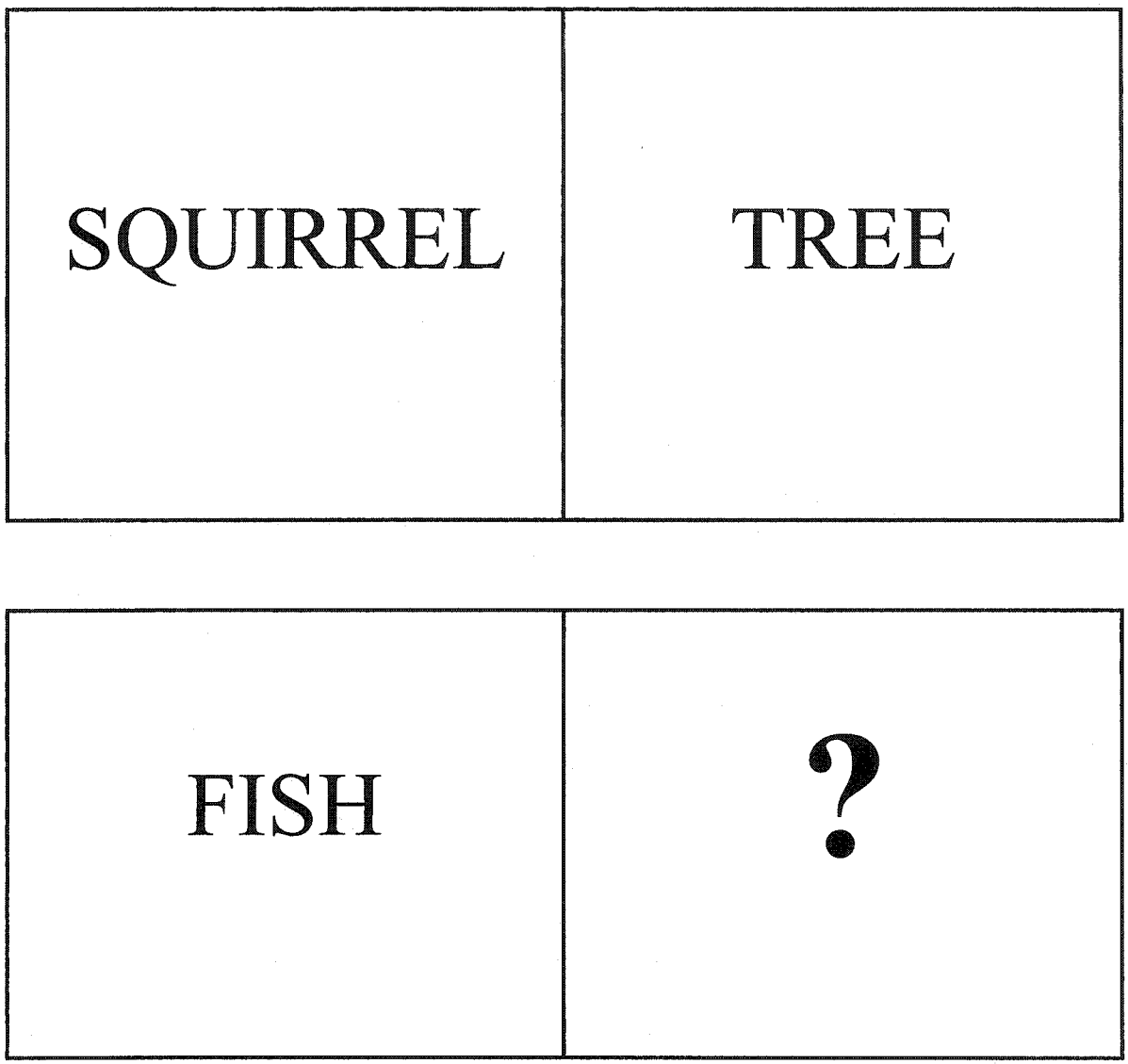
GAT (Version 1) Set B Question 6



GAT (Version 1) Set B Question 7

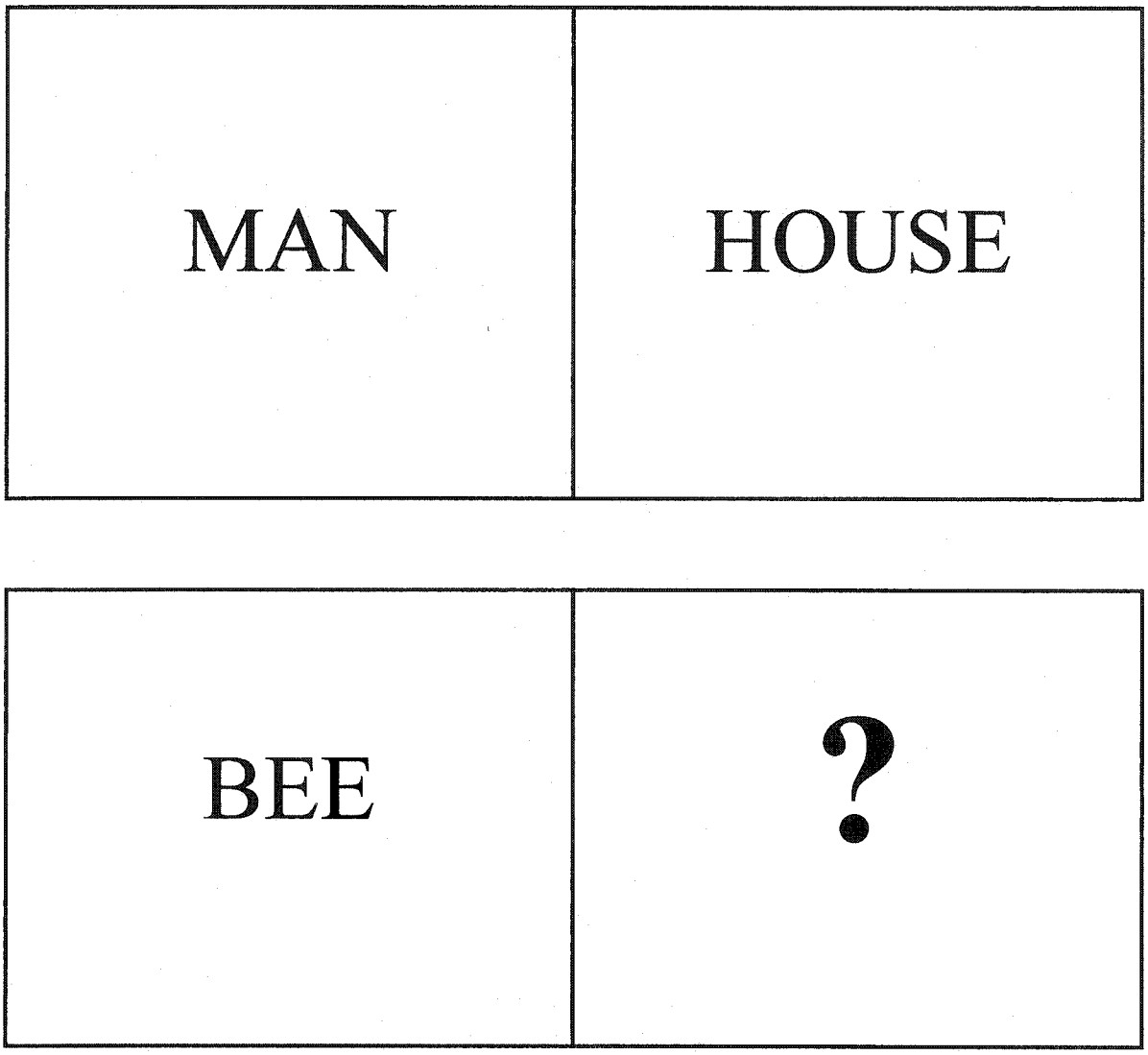


GAT (Version 1) Set B Question 8



A	B	C	D	E
SHARK	LIZARD	ROSE	OCEAN	HOOK

GAT (Version 1) Set B Question 9



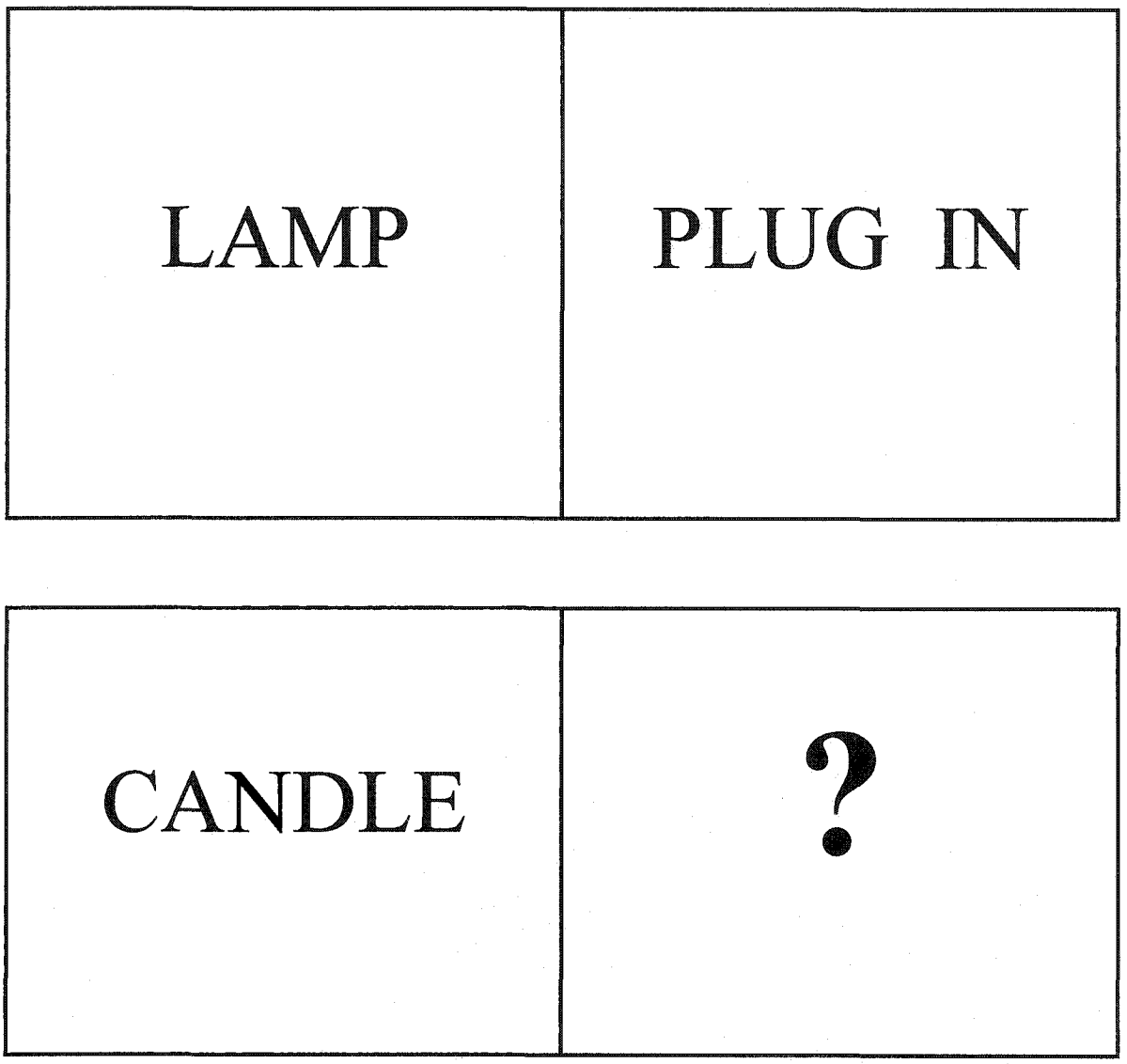
GRAPES	HIVE	HONEY	BUTTERFLY	FLY
A	B	C	D	E

GAT (Version 1) Set B Question 10

TOOTHBRUSH	TEETH
SOAP	?

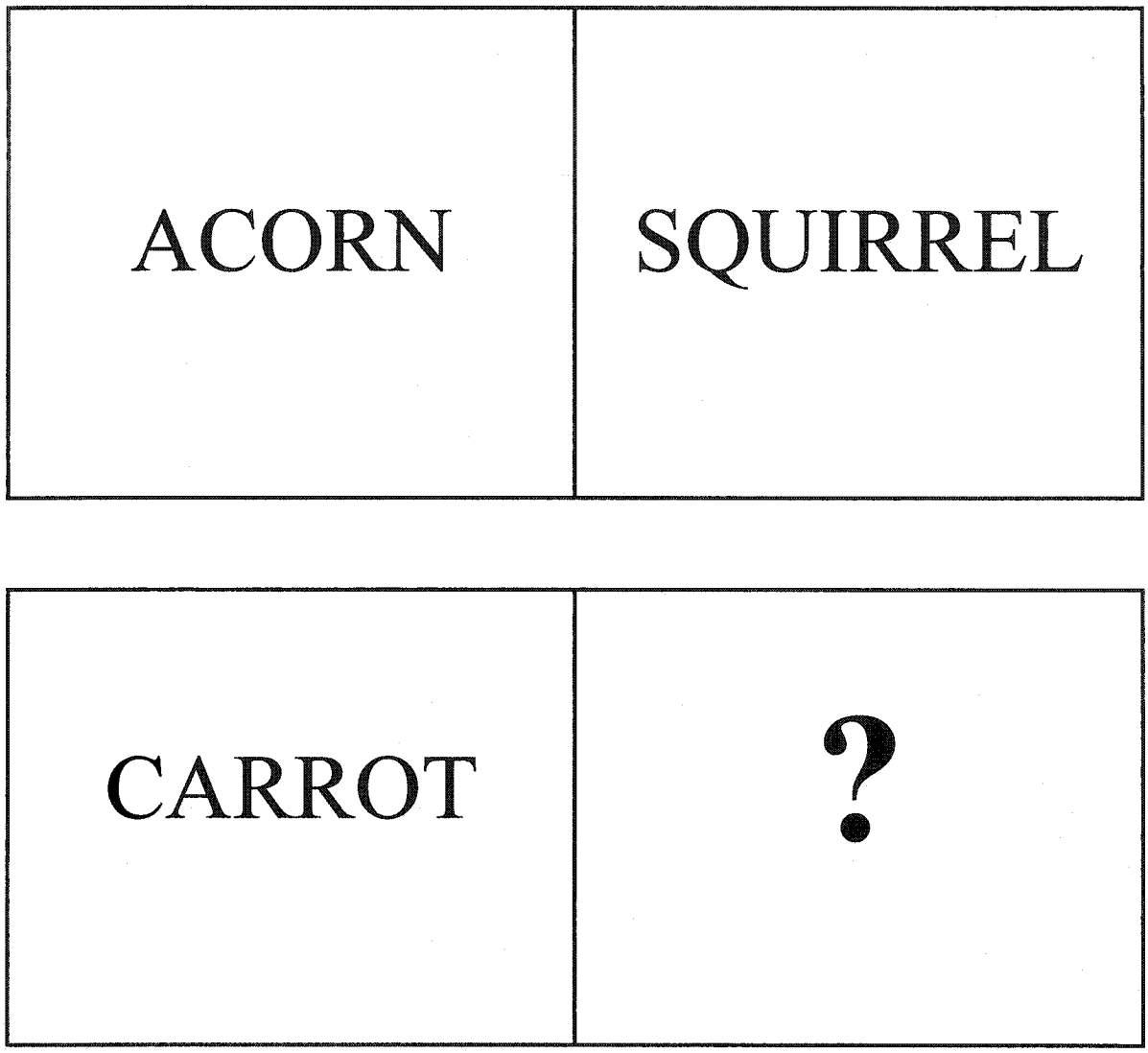
A	B	C	D	E
TOWEL	ICEBERG	COUCH	HANDS	WATER

GAT (Version 1) Set B Question 11



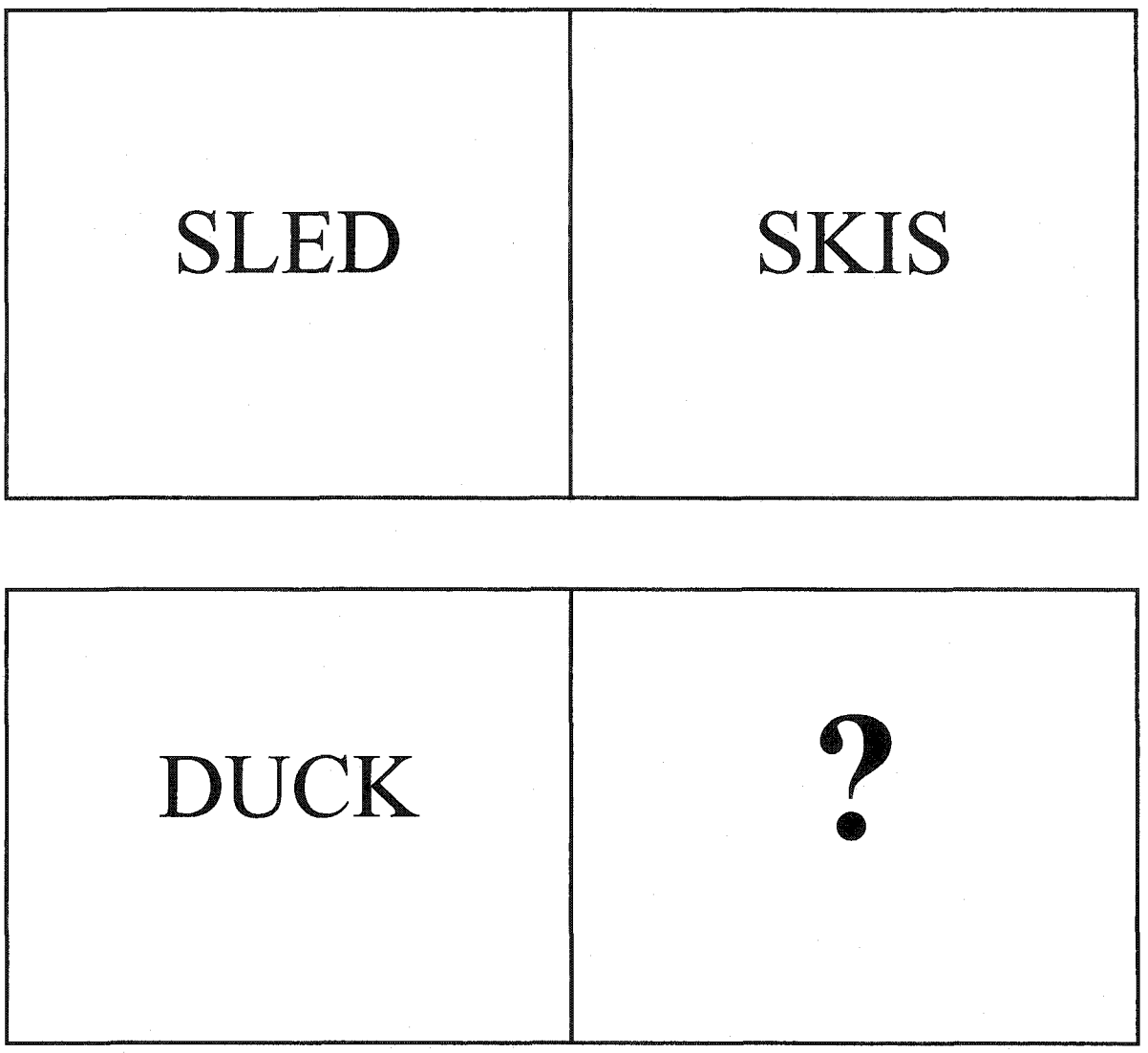
A	B	C	D	E
MATCHES	FIRE	FLASHLIGHT	LANTERN	LETTER

GAT (Version 1) Set B Question 12



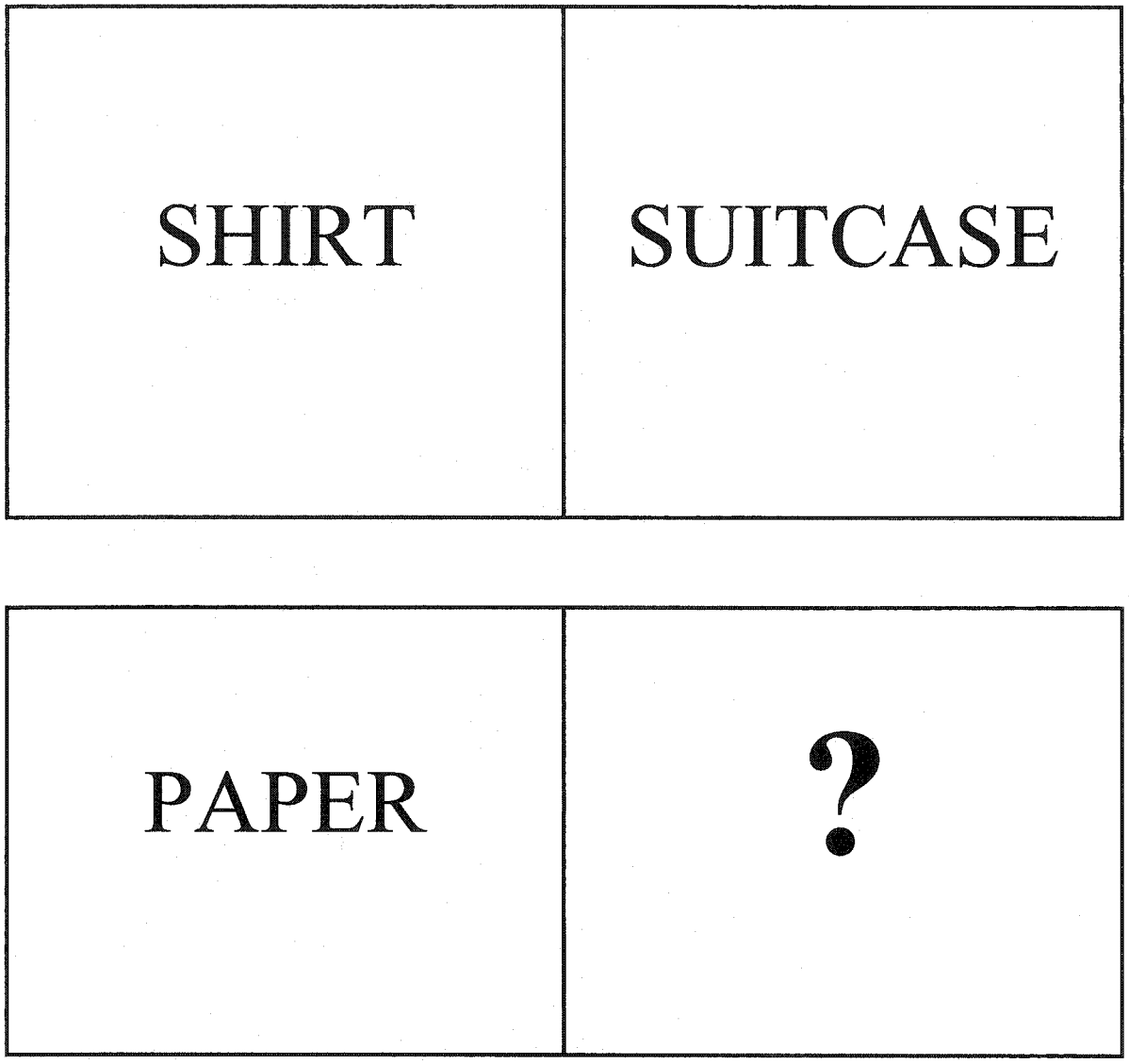
A	B	C	D	E
CELERY	PAINTBRUSH	LIGHT BULB	MUSHROOMS	BUNNY

GAT (Version 1) Set B Question 13



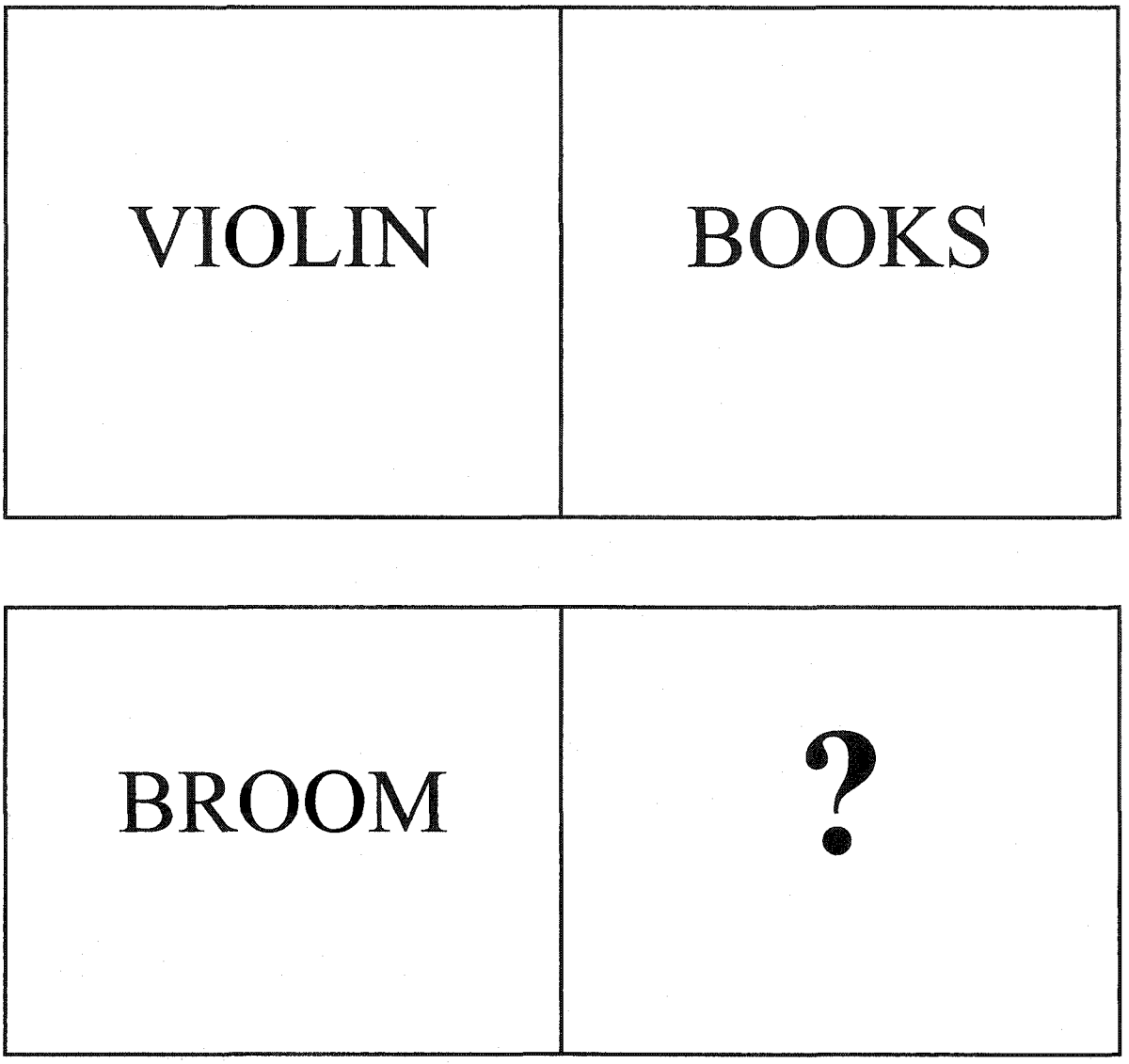
A	B	C	D	E
LIFEJACKET	WATER	SAXOPHONE	PENNY	PIG

GAT (Version 1) Set B Question 14



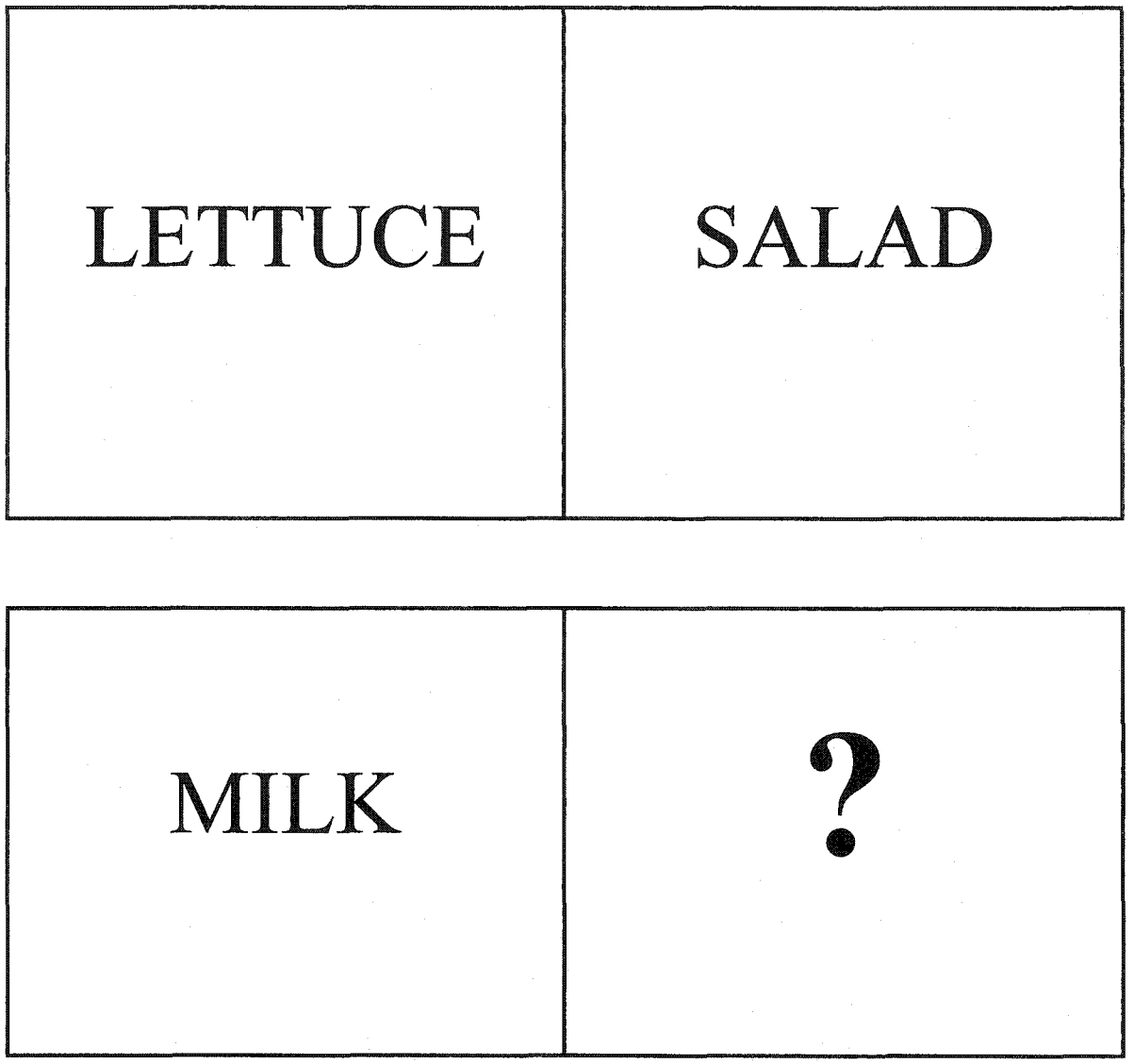
BRIEFCASE	PENCIL	DRAPE	CHAIR	TAPE
A	B	C	D	E

GAT (Version 1) Set B Question 15

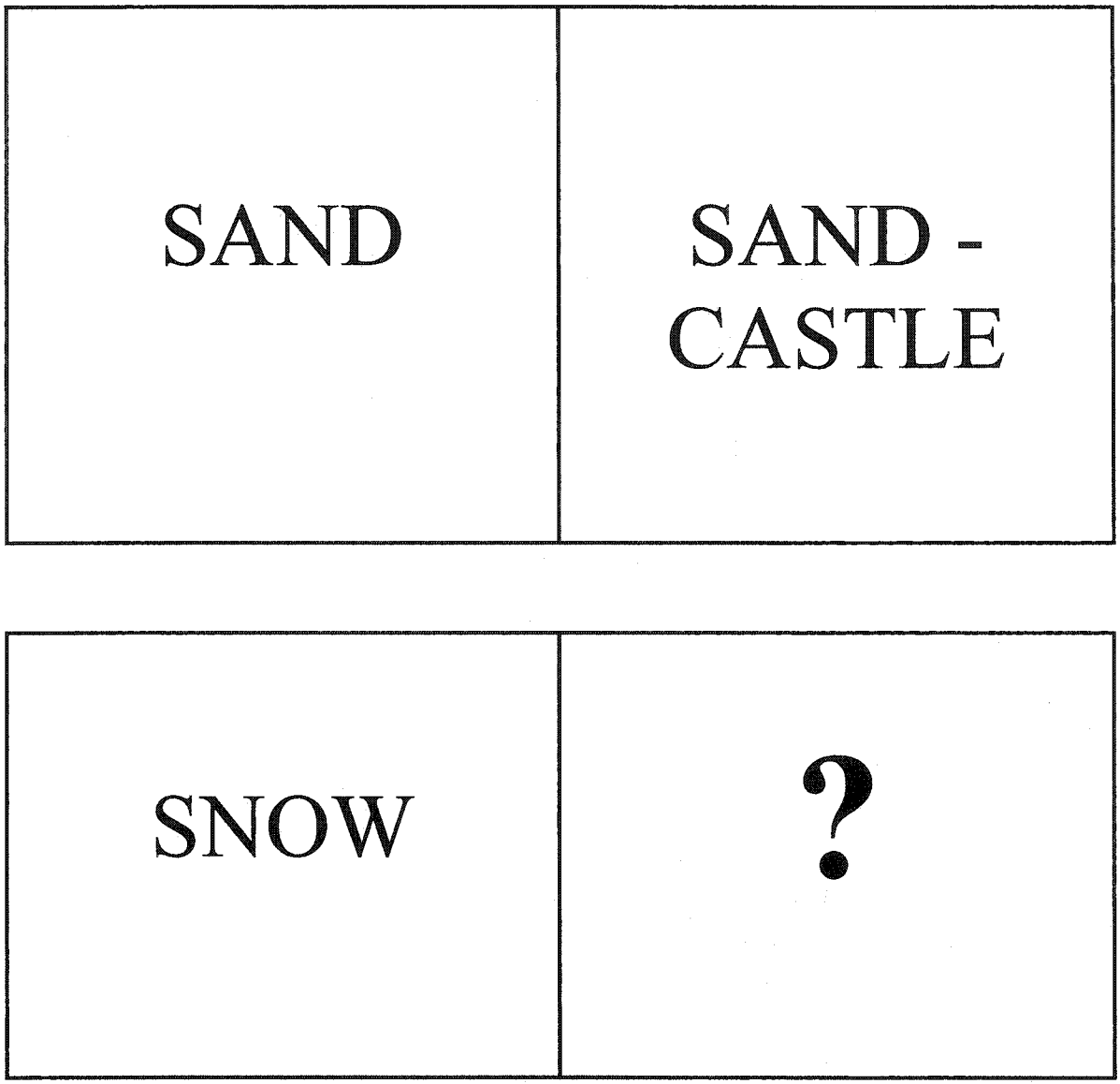


A	B	C	D	E
PORCUPINE	MAID	TREE	DICE	WITCH

GAT (Version 1) Set B Question 16

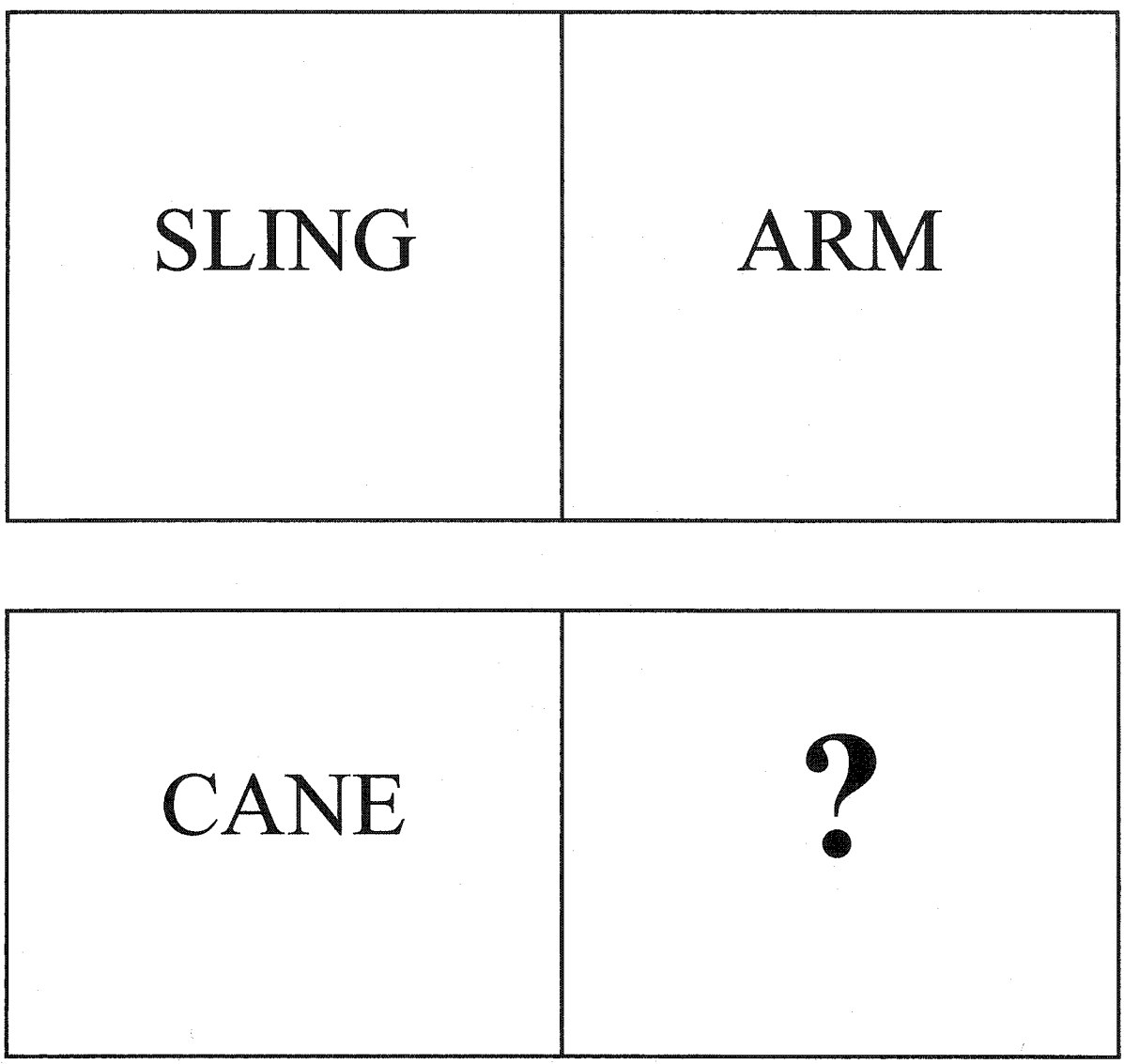


GAT (Version 1) Set B Question 17



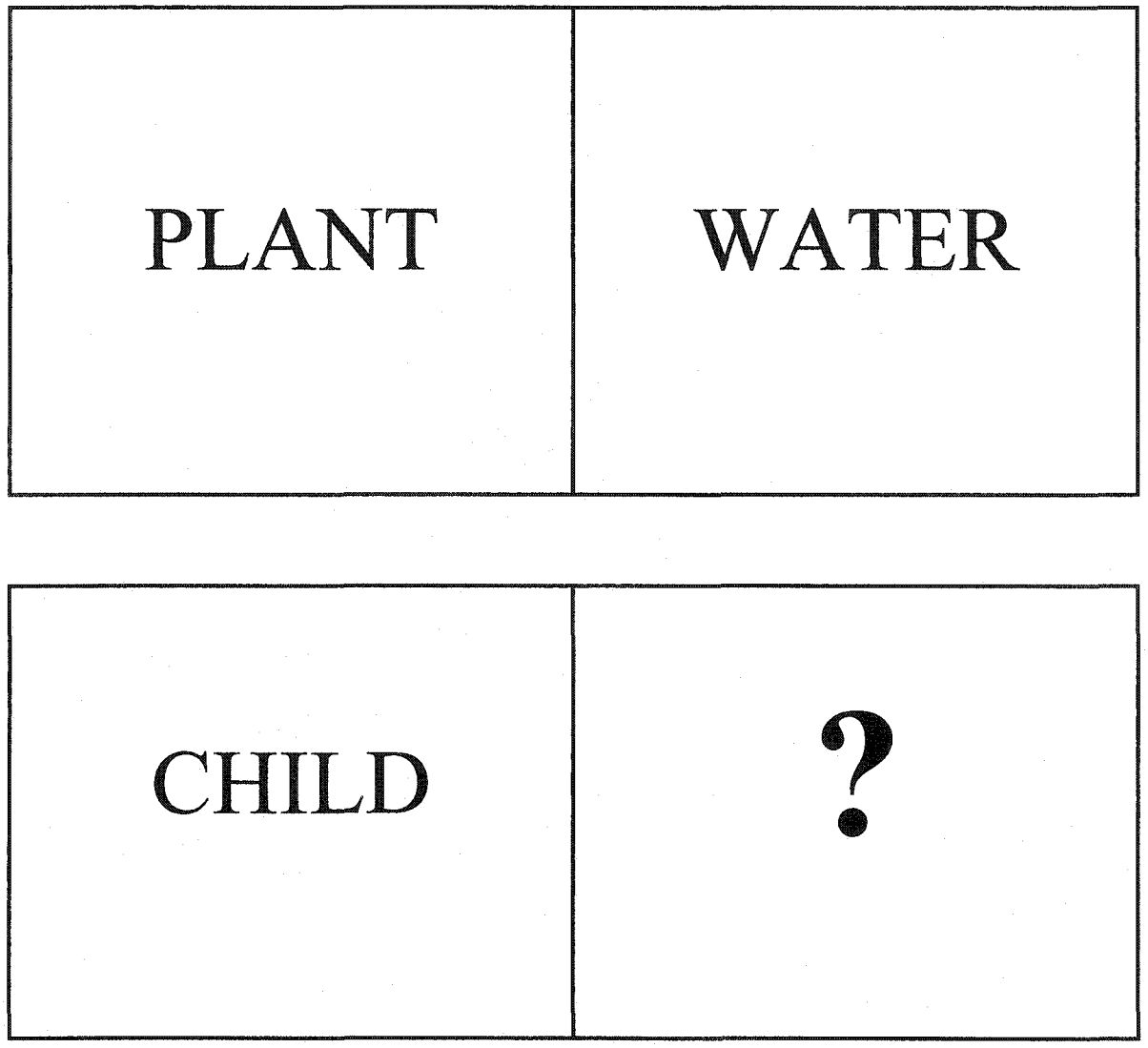
BALLOON	RAIN	SNOWMAN	WINTER	ICE
A	B	C	D	E

GAT (Version 1) Set B Question 18



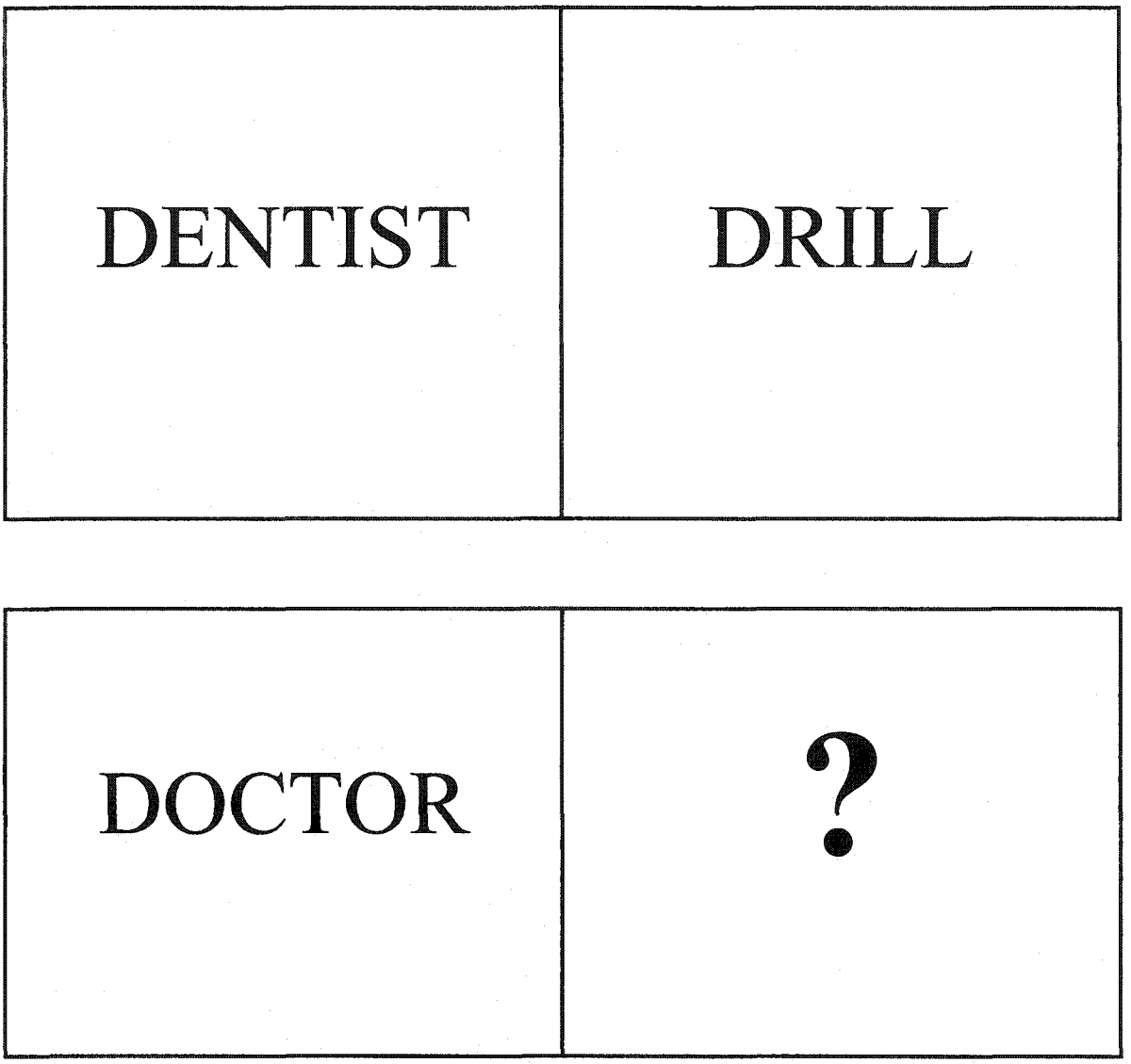
A	B	C	D	E
GRANDMA	CANDY-CANE	LADDER	WALKER	LEG

GAT (Version 1) Set B Question 19



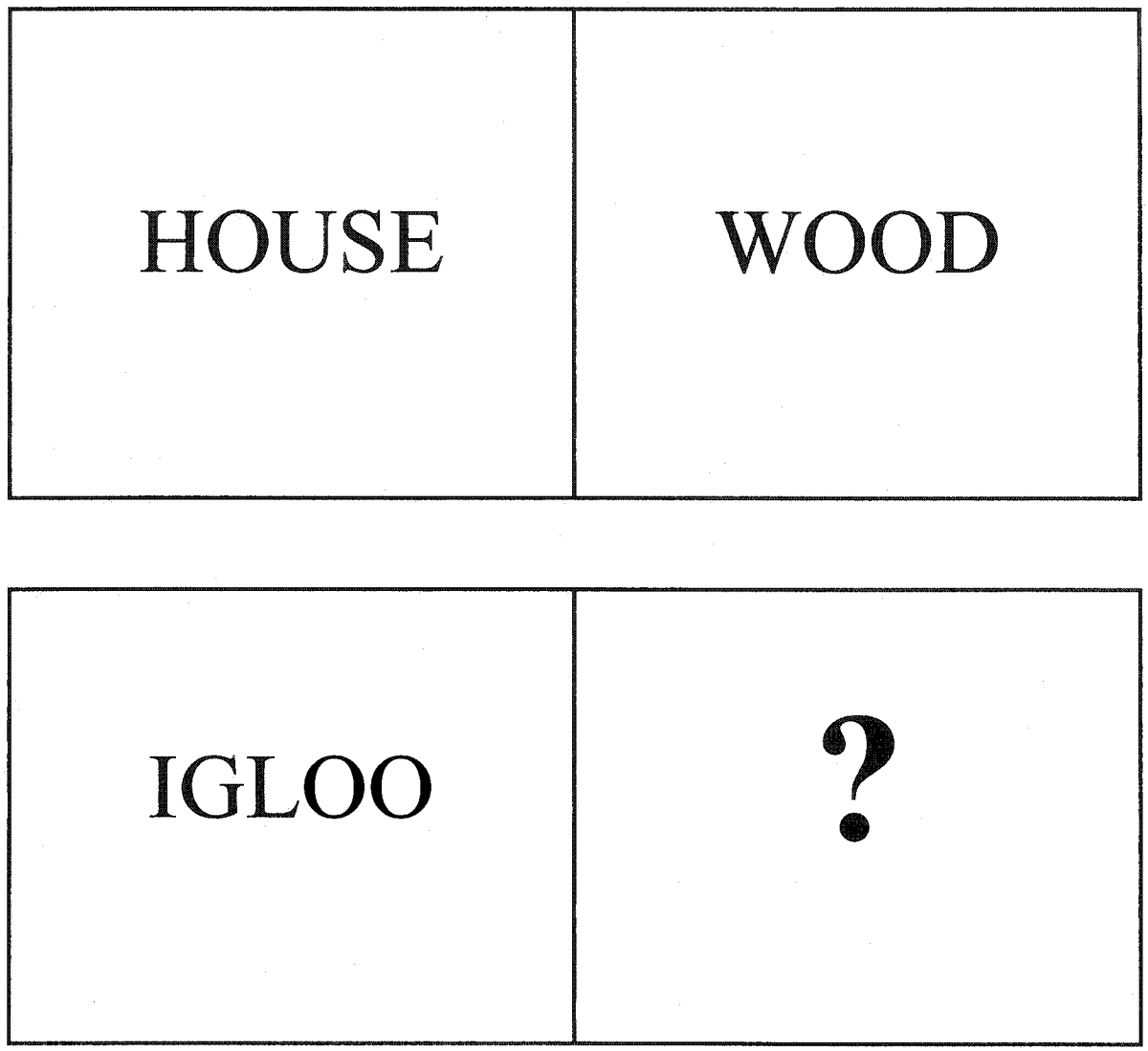
FOOD	MOTHER	TODDLER	SCALE	BABY
A	B	C	D	E

GAT (Version 1) Set B Question 20



A	B	C	D	E
SURGICAL MASK	STETHOSCOPE	NURSE	LAB TECHNICIAN	VACUUM

GAT (Version 1) Set B Question 21



SNOW	PAN	TEEPEE	KETTLE	ESKIMO
A	B	C	D	E

GAT (Version 1) Set B Question 22

RAKE	YARD WORK
SCISSORS	?

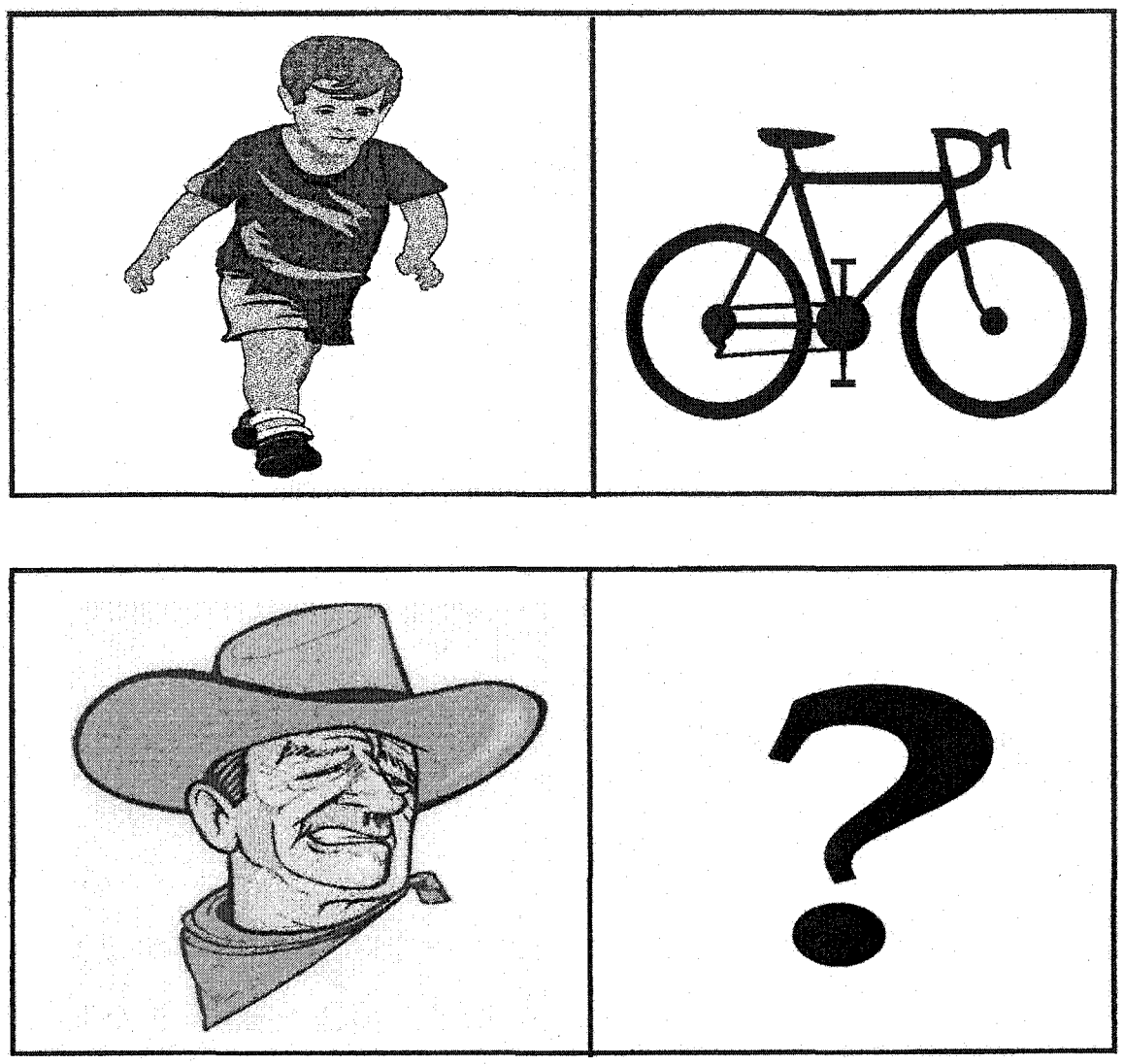
A	B	C	D	E
TAPE	SEWING	PAPER	PLIERS	GLASSES

GAT (Version 1) Set B Question 23

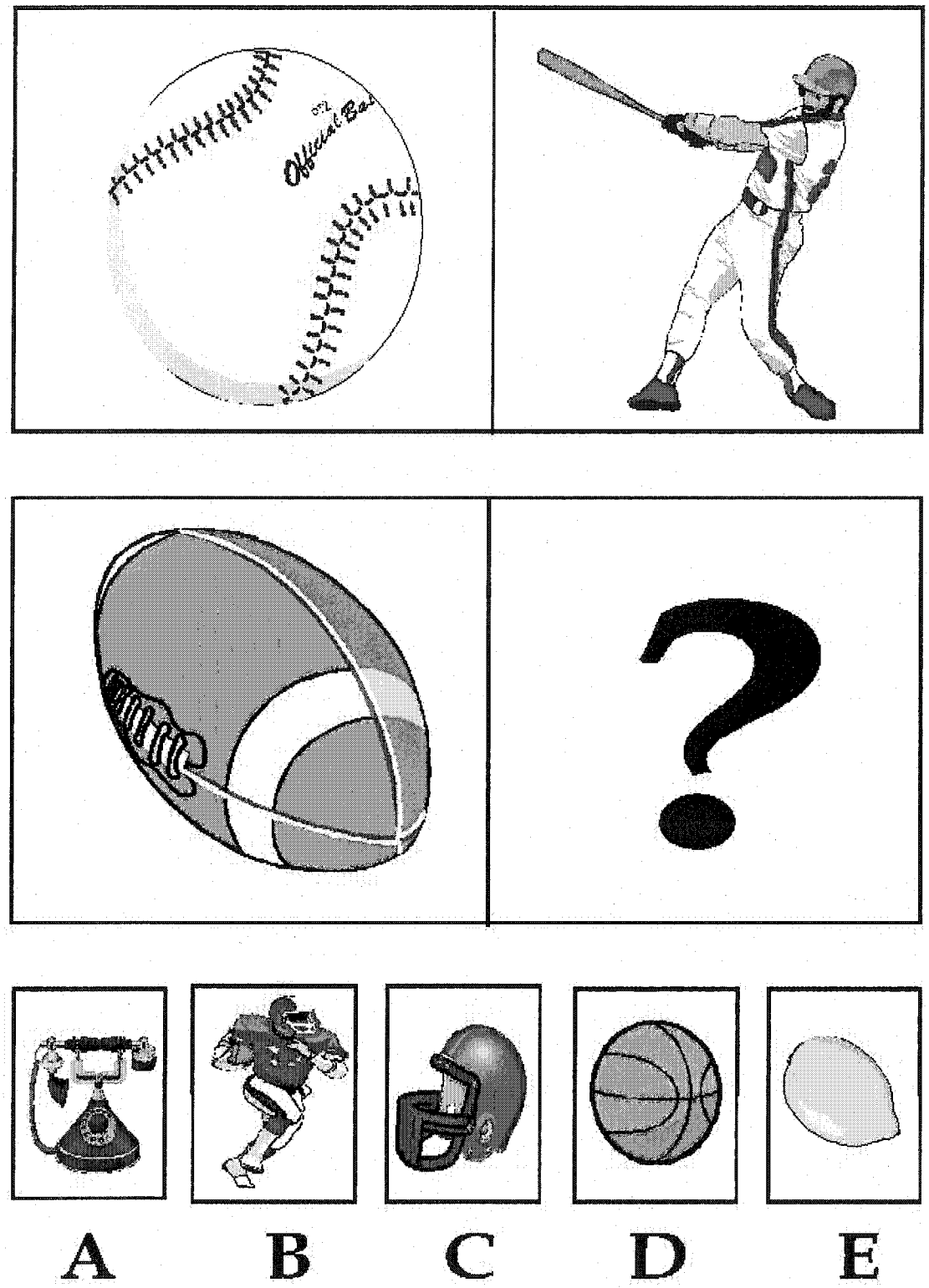
SANDWICH	COOLER
HAMMER	?

A	B	C	D	E
GAVEL	BOW	SAW	TOOLBOX	NAIL

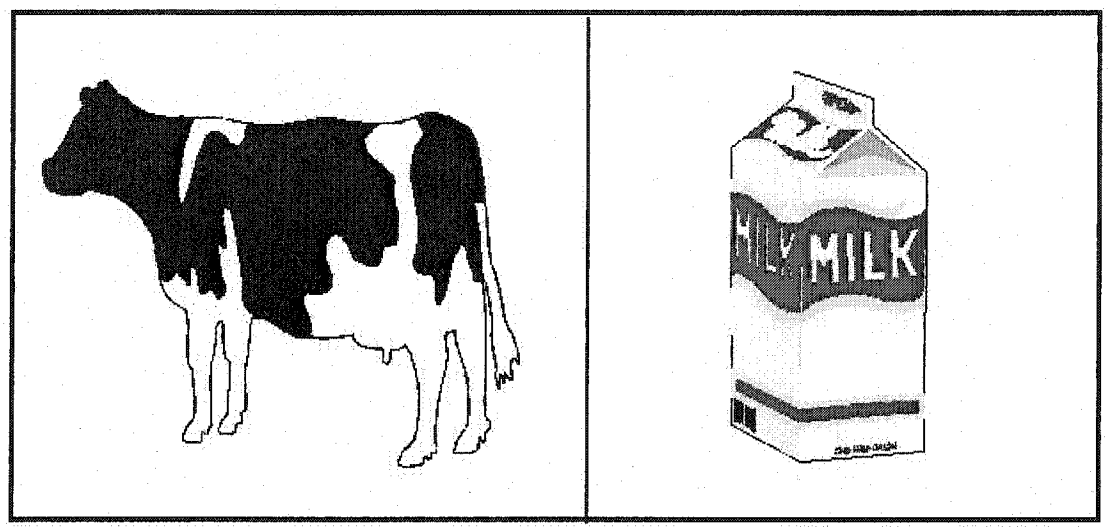
GAT (Version 1) Set B Question 24

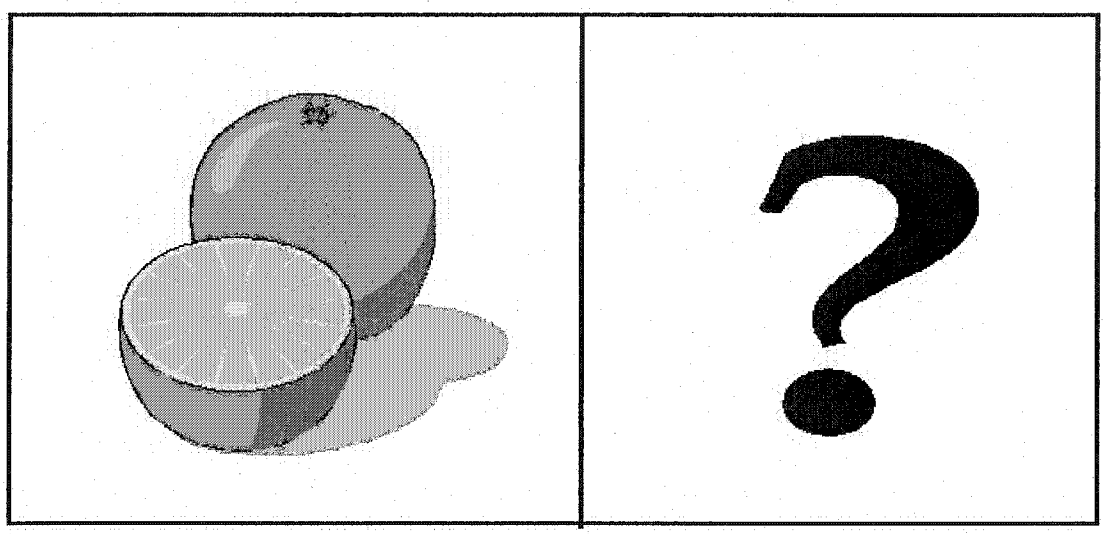


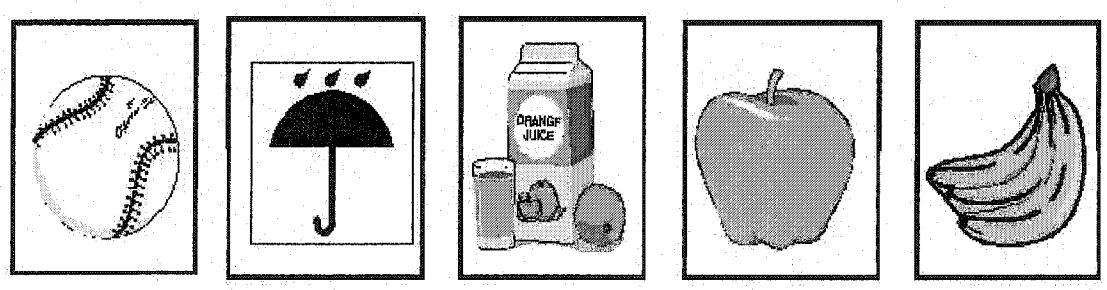
GAT (Version 1) Set A Question 1



GAT (Version 1) Set A Question 2

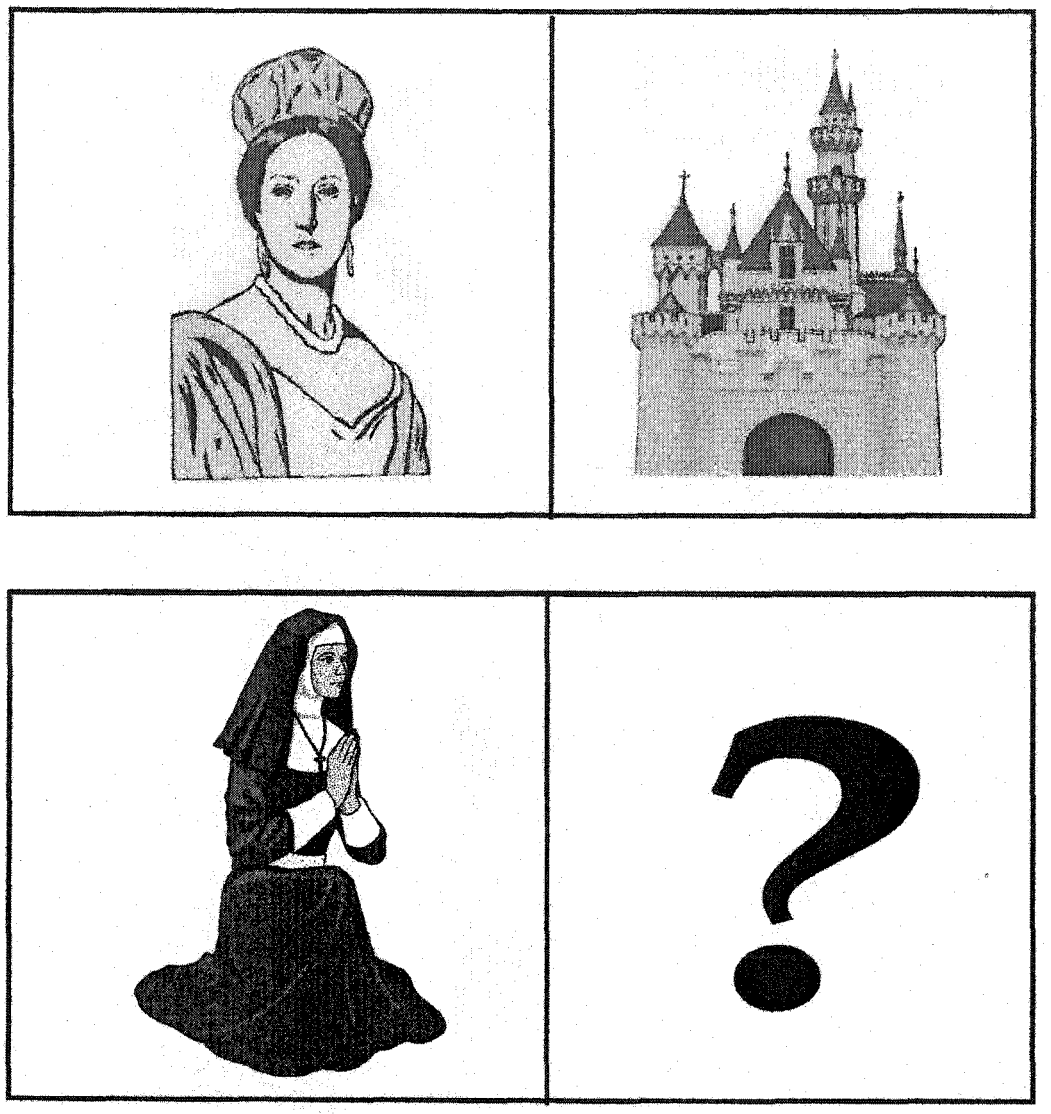




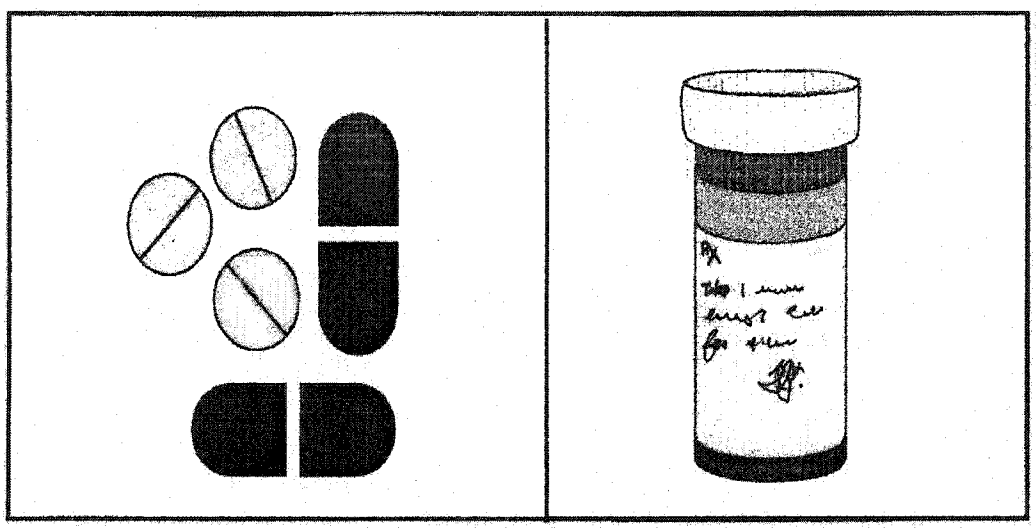


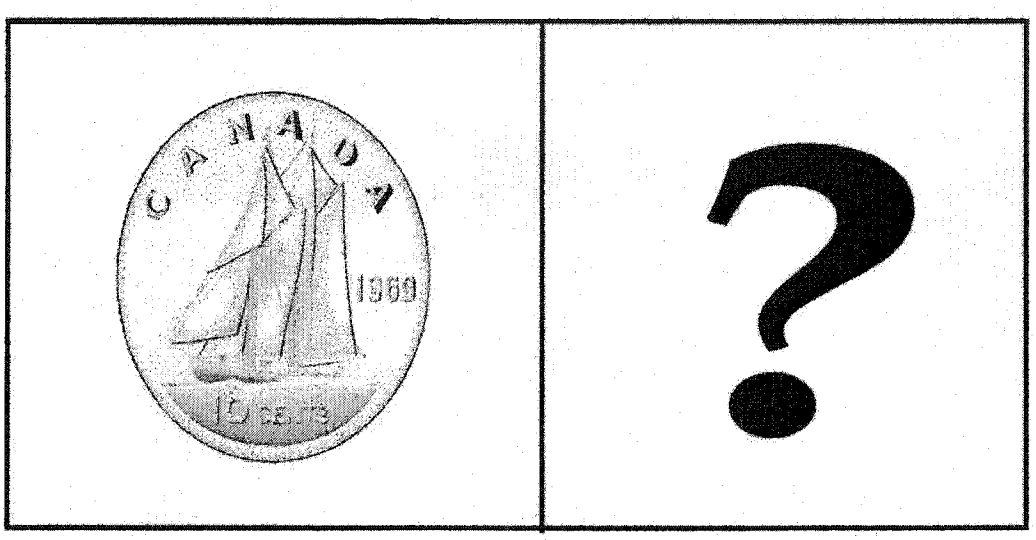
A B C D E

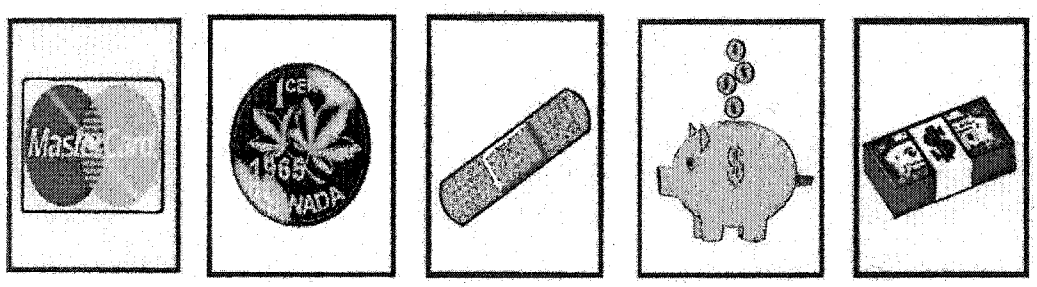
GAT (Version 1) Set A Question 3



GAT (Version 1) Set A Question 4

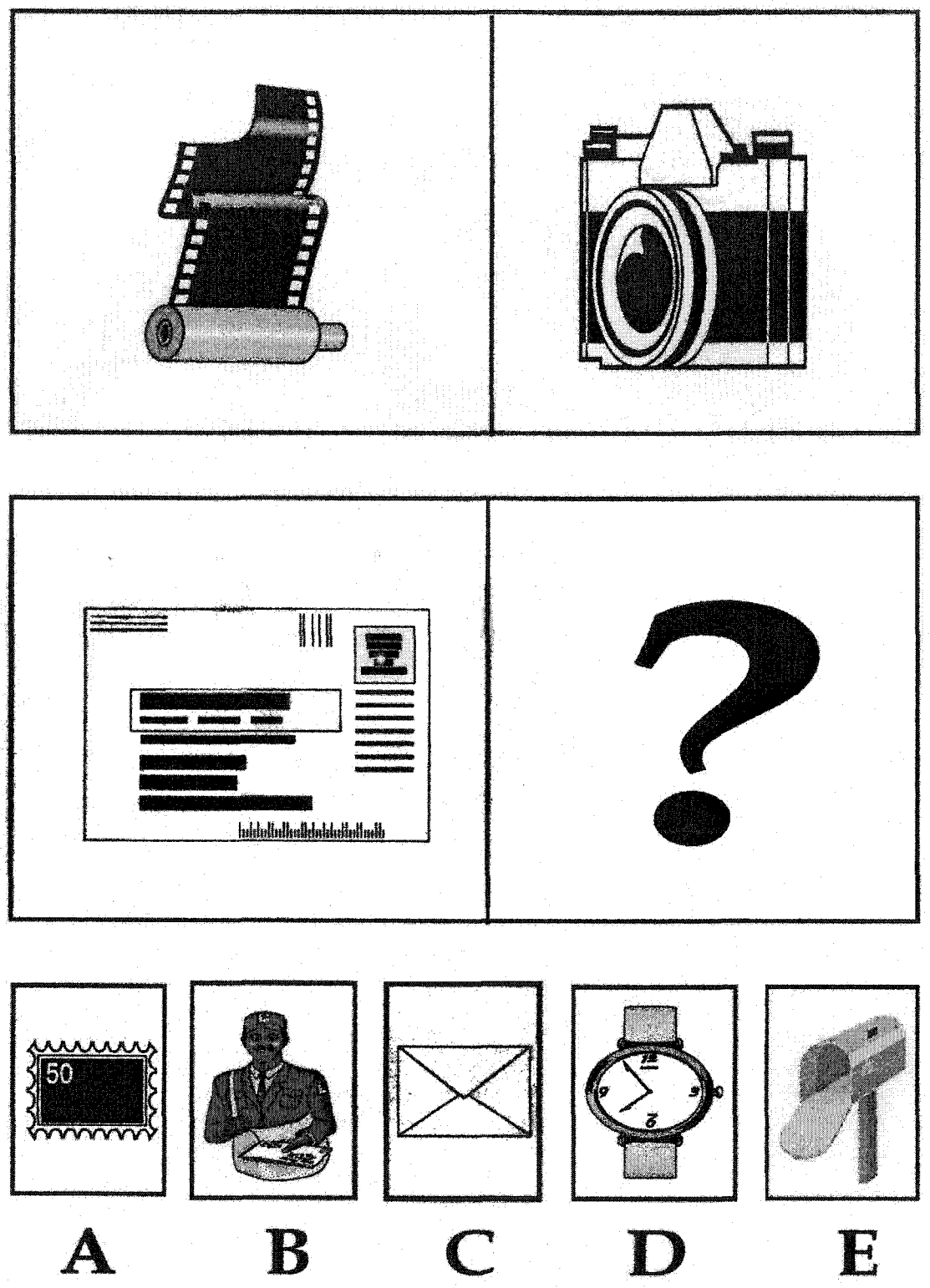




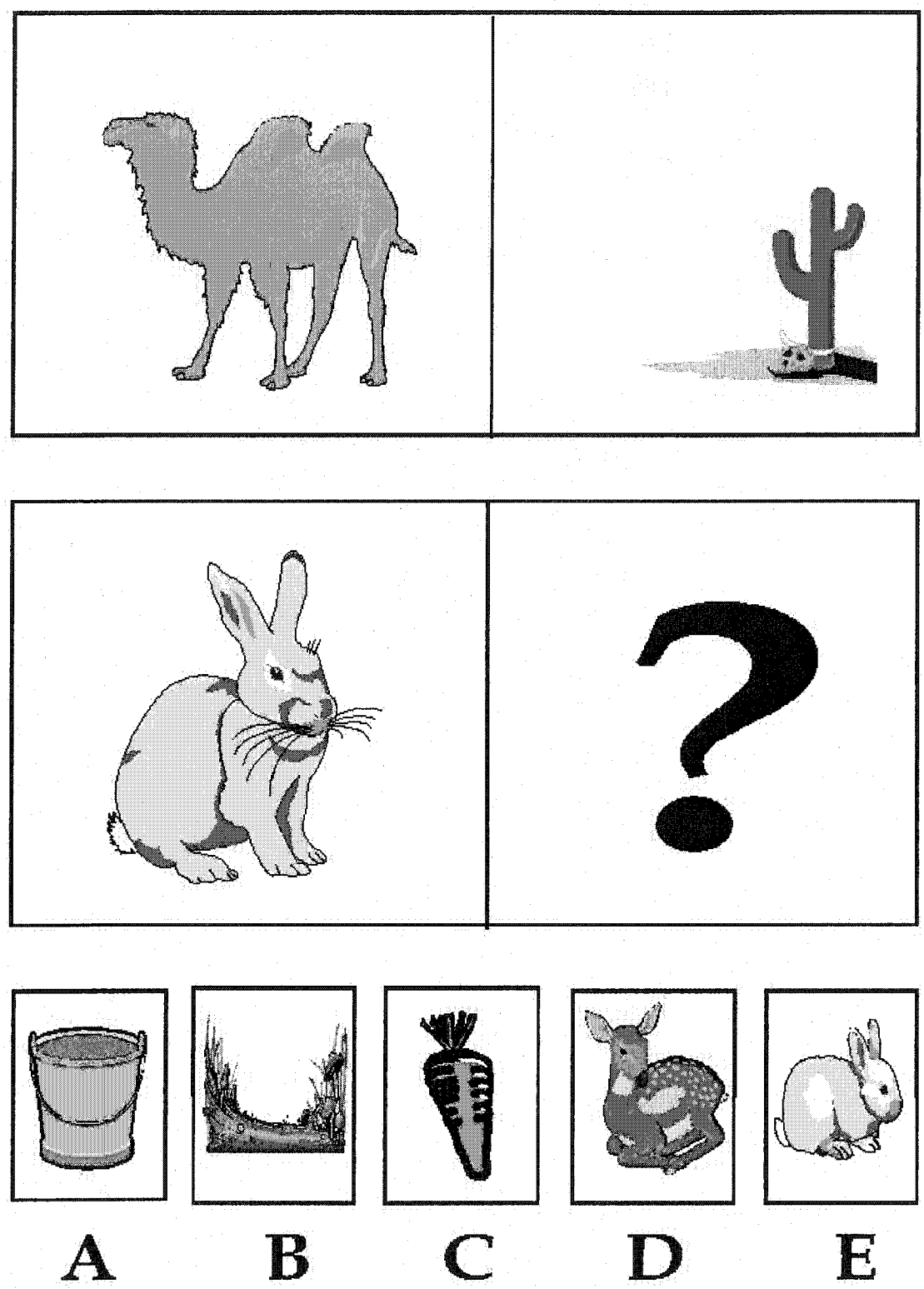


A B C D E

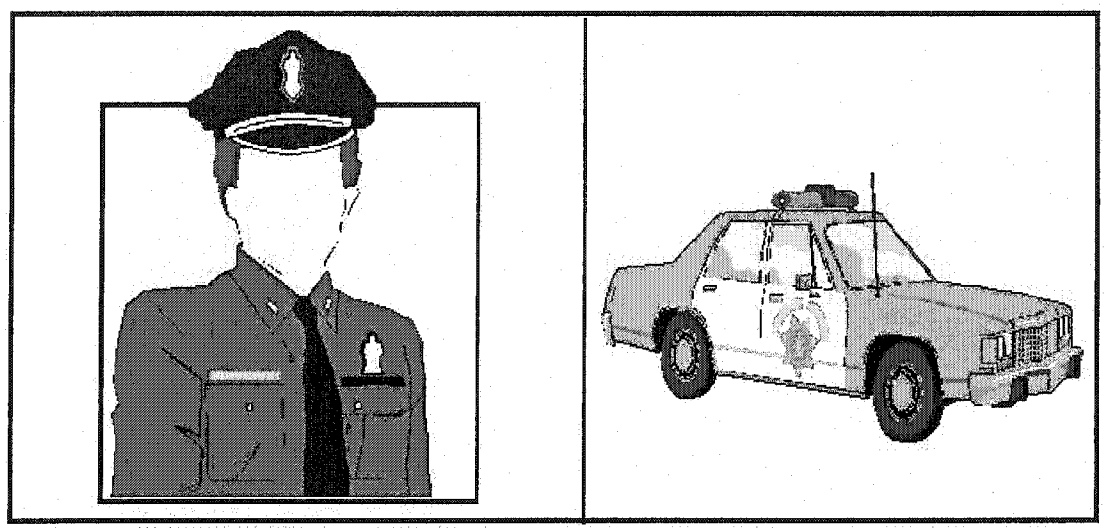
GAT (Version 1) Set A Question 5

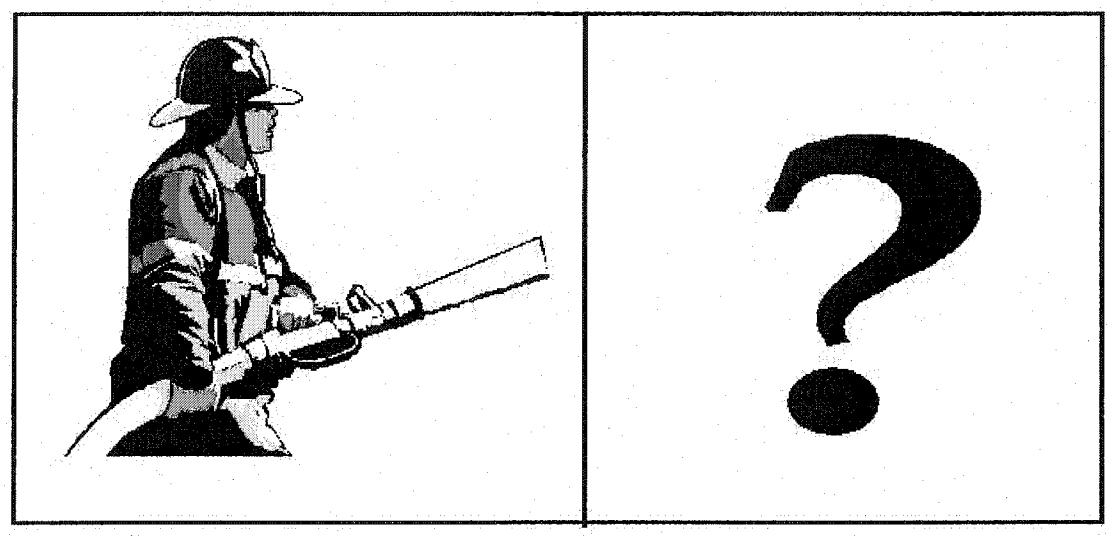


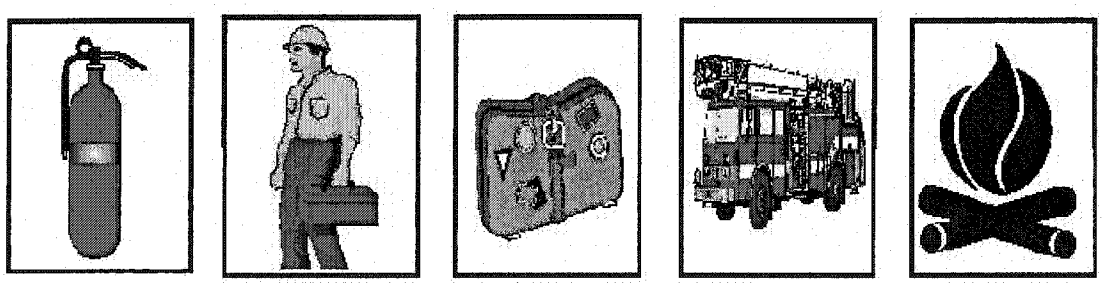
GAT (Version 1) Set A Question 6



GAT (Version 1) Set A Question 7

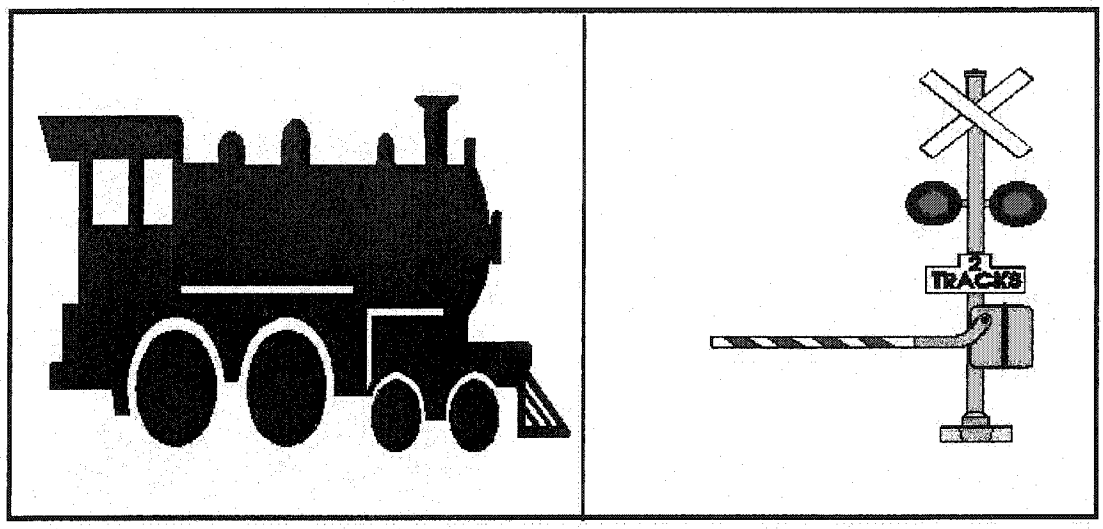


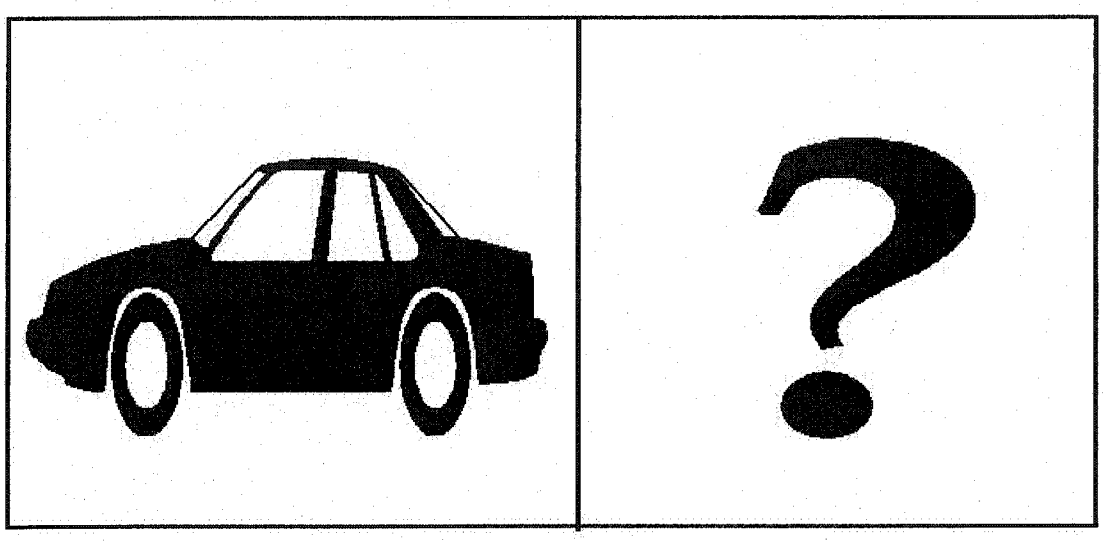


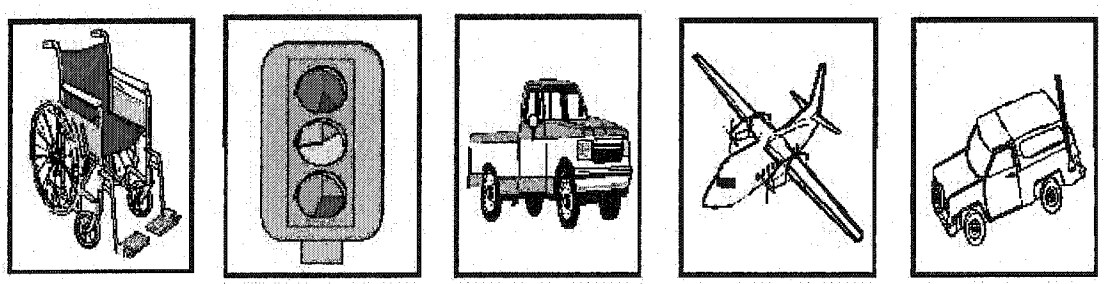


A B C D E

GAT (Version 1) Set A Question 8

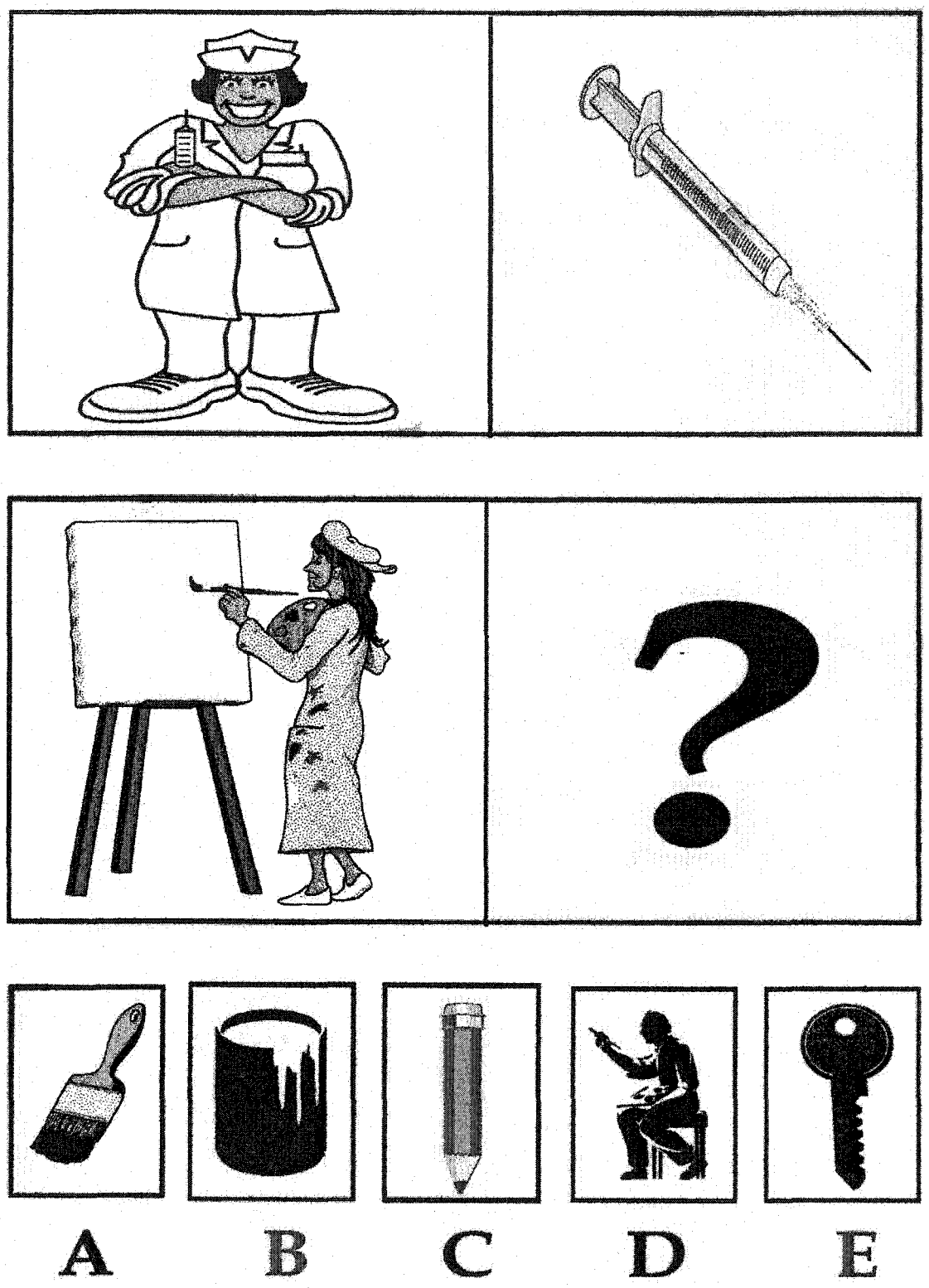




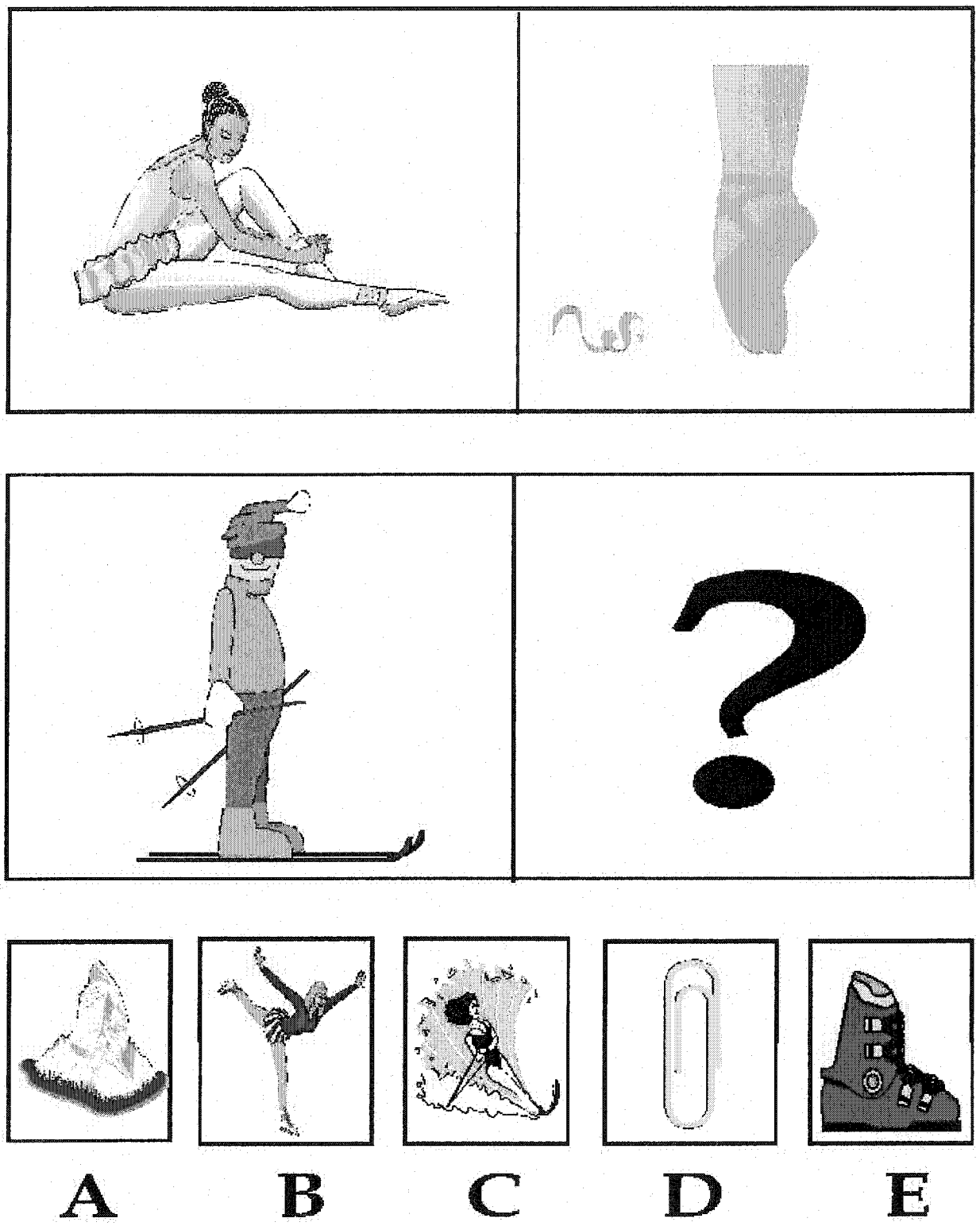


A B C D E

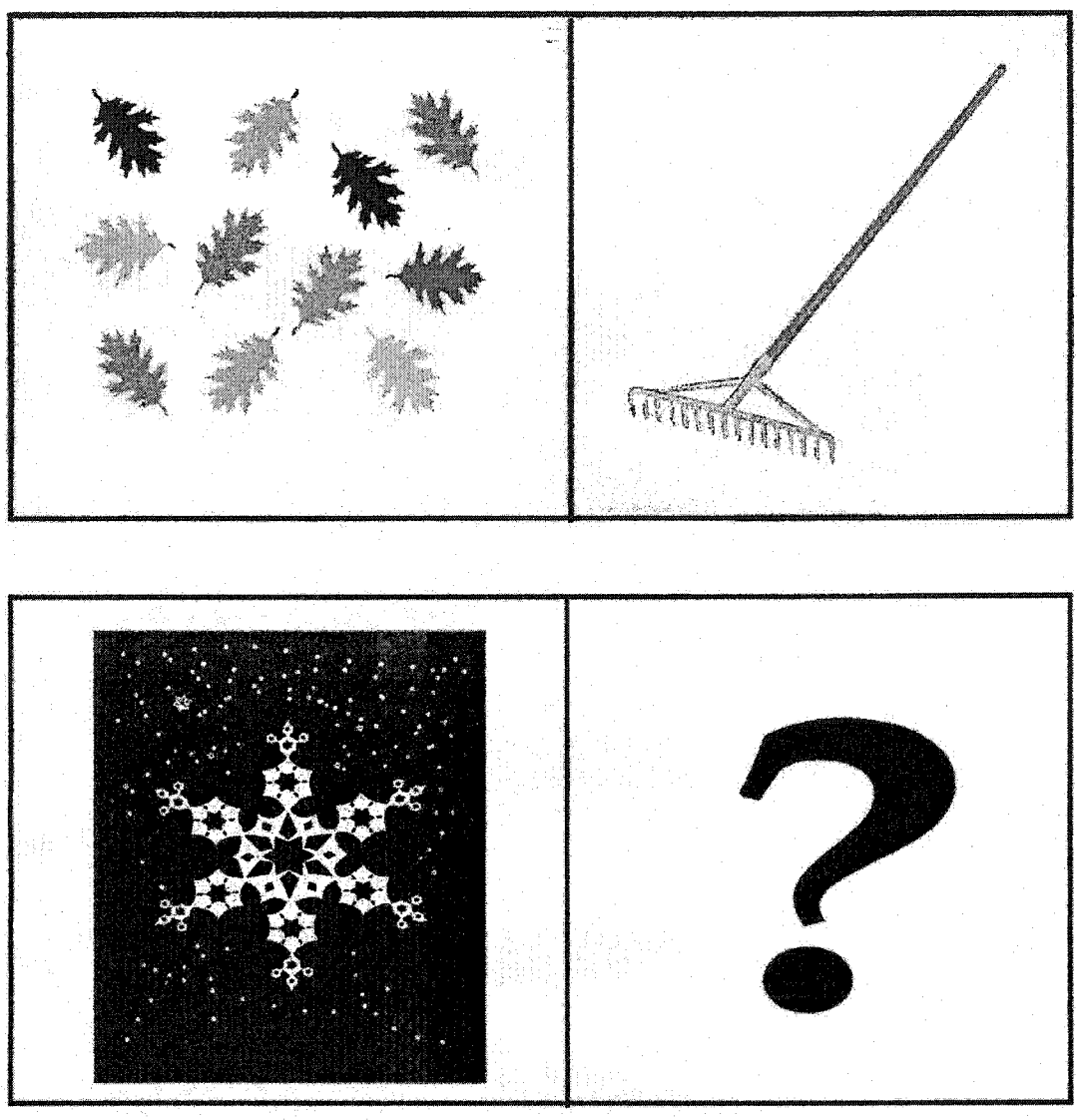
GAT (Version 1) Set A Question A9



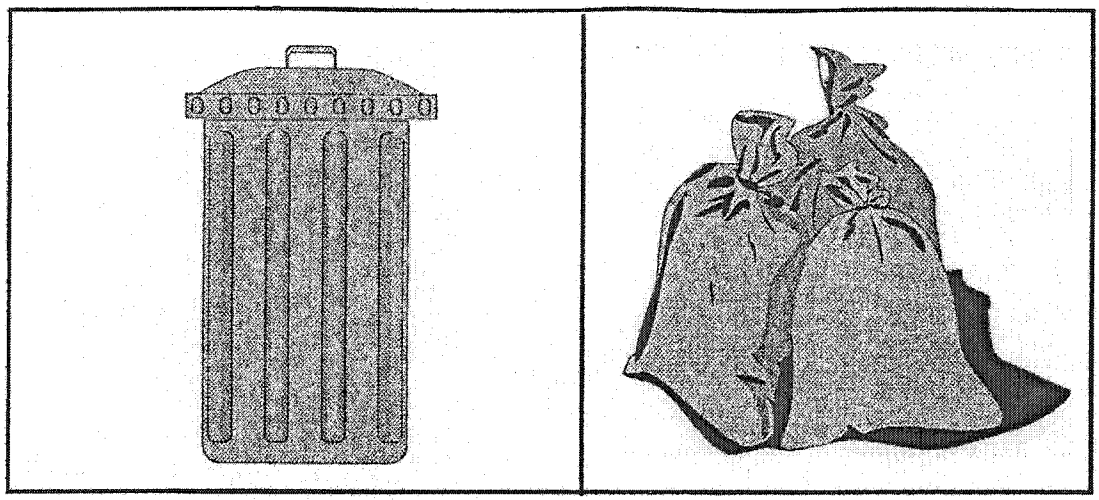
GAT (Version 1) Set A Question 10

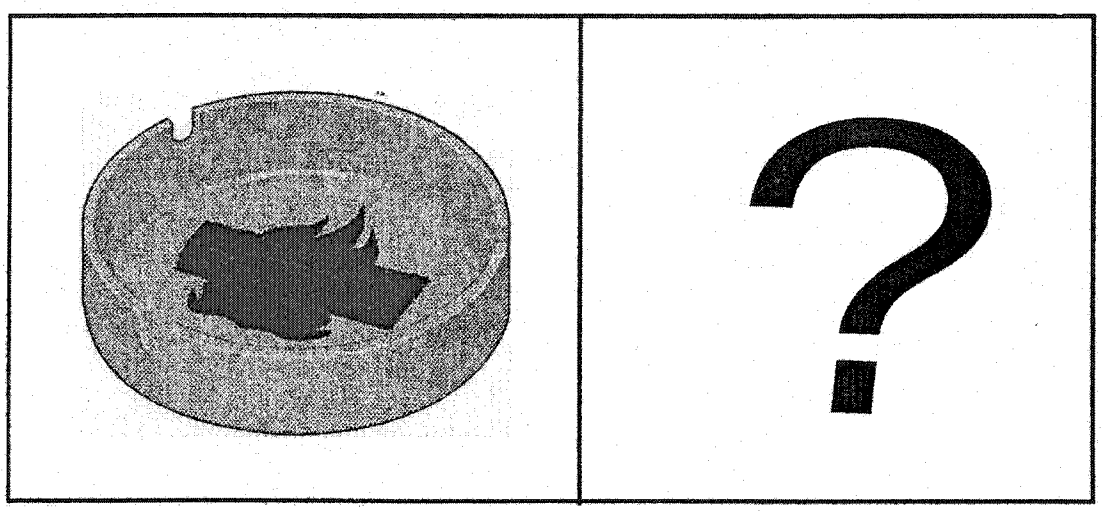


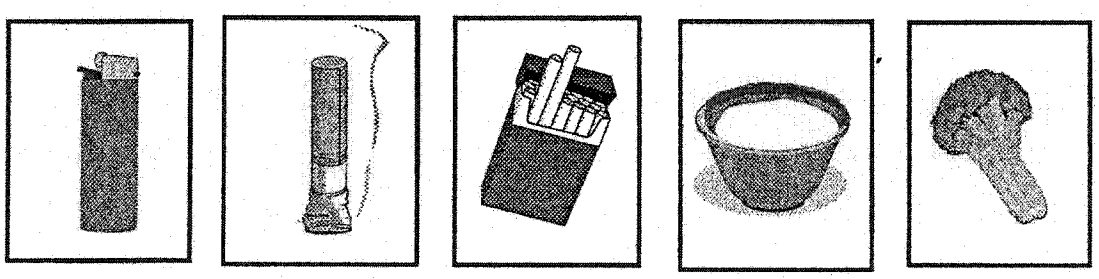
GAT(Version 1) Set A Question 11



GAT (Version 1) Set A Question 12

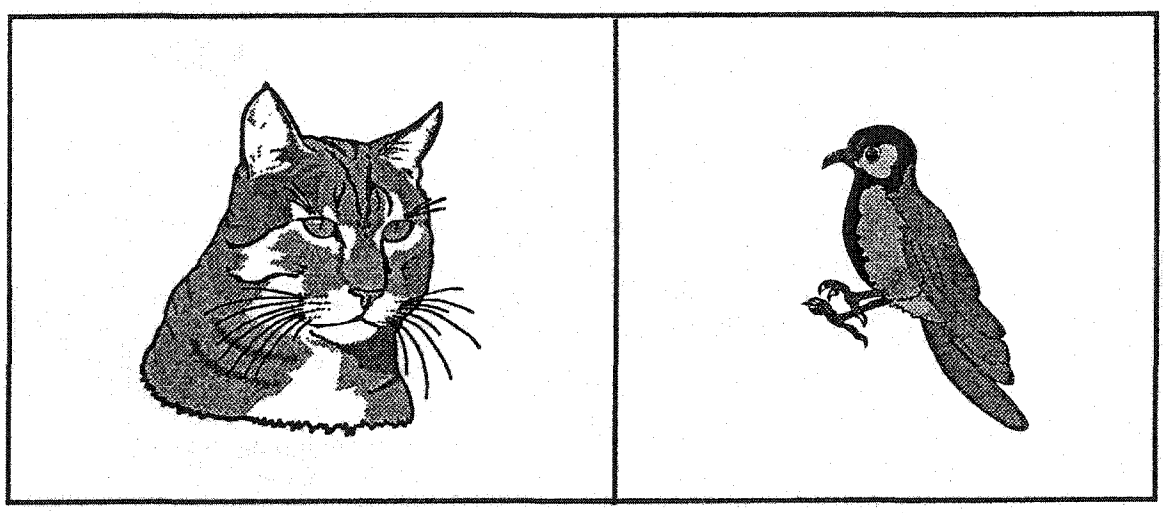


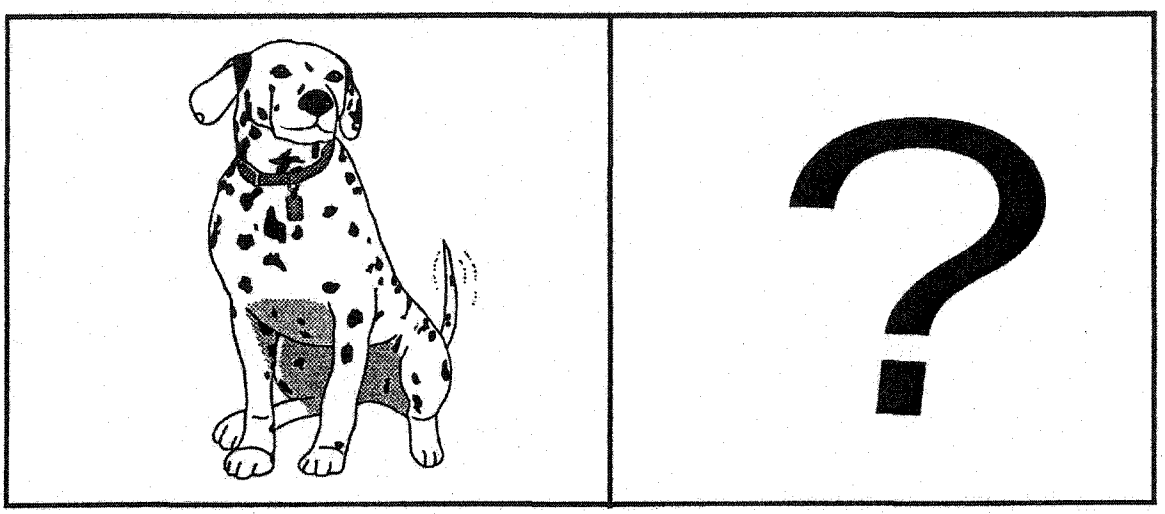


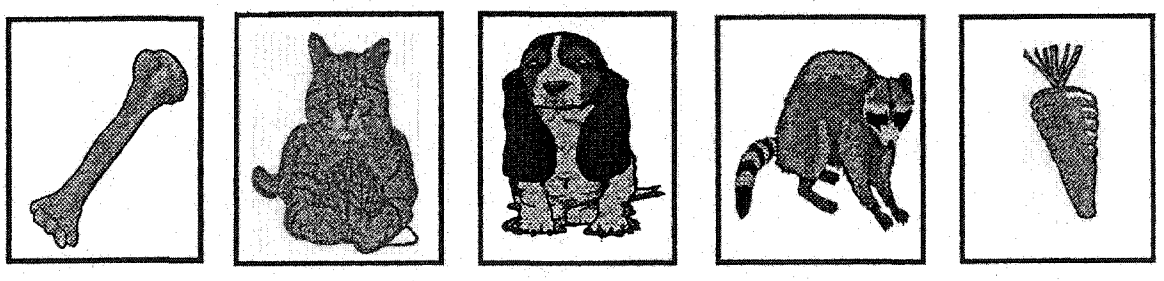


A B C D E

GAT (Version 1) Set A Question 13

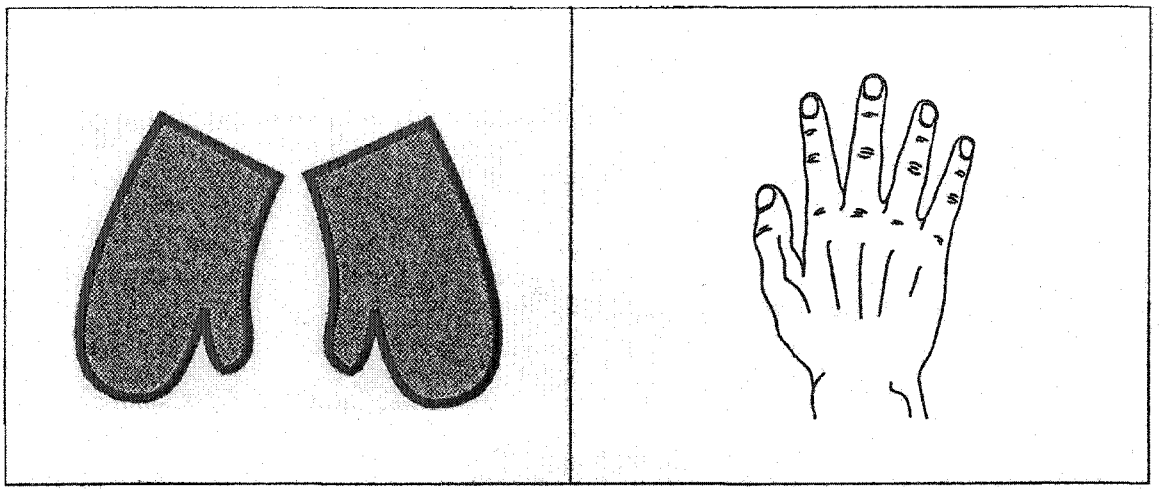


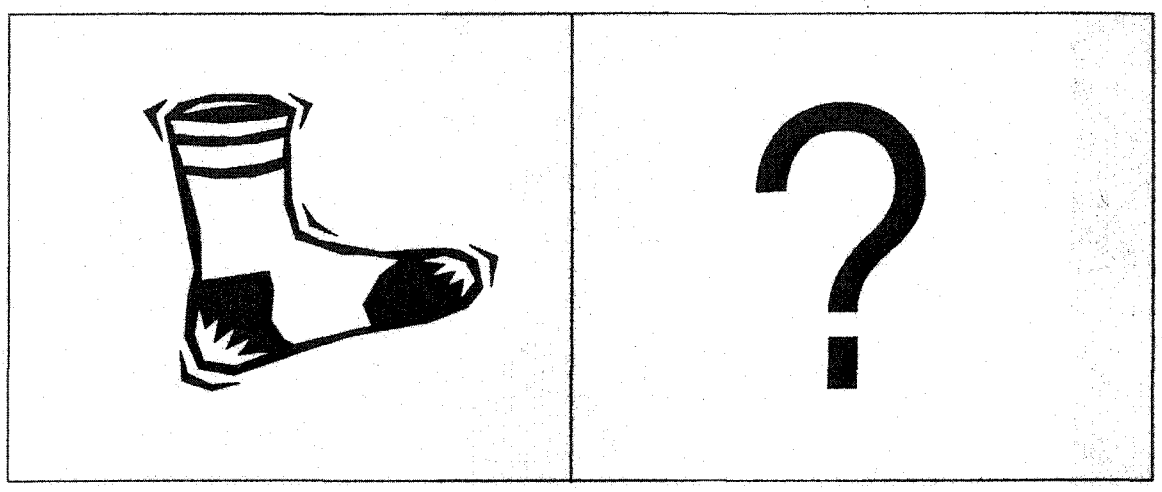


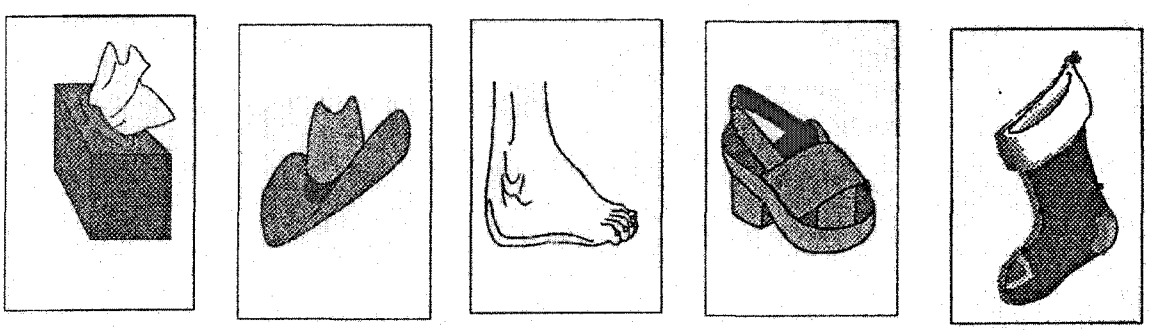


A B C D E

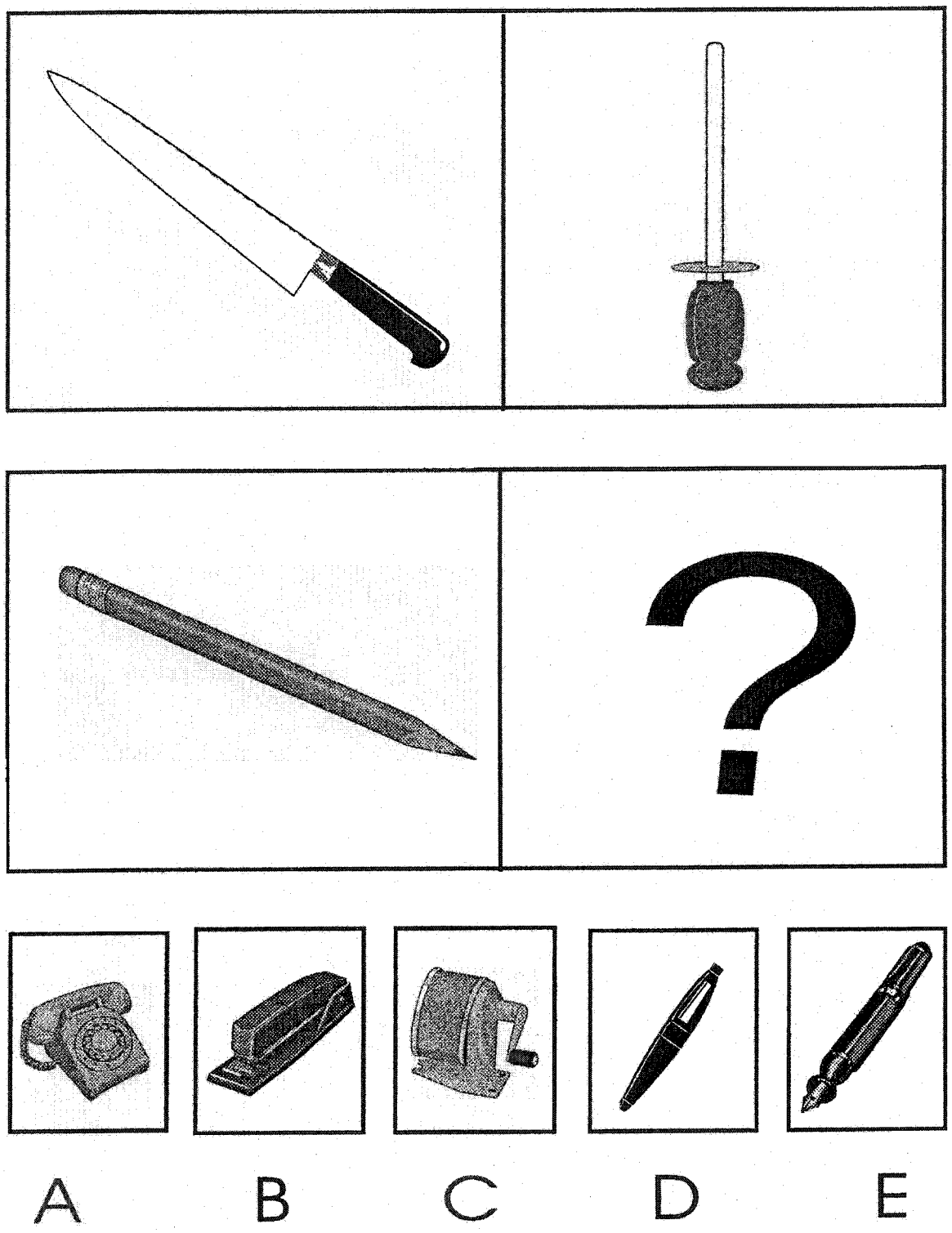
GAT (Version 1) Set A Question 14



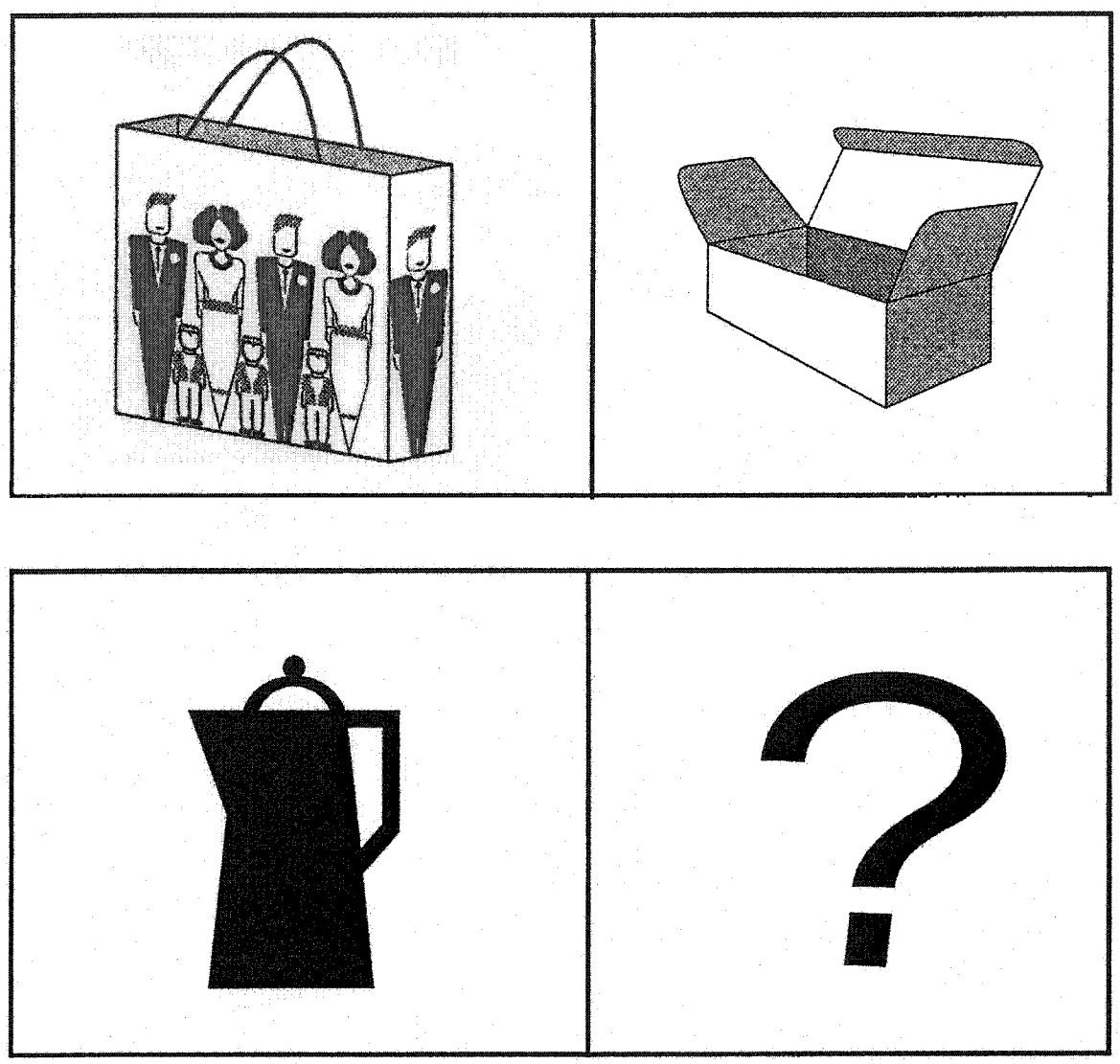




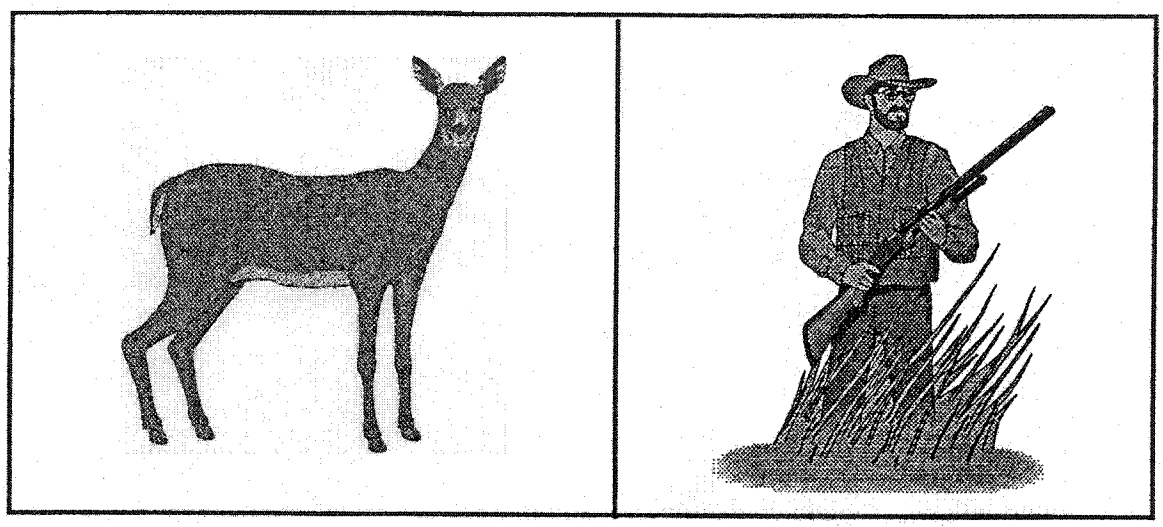
GAT (Version 1) Set A Question 15

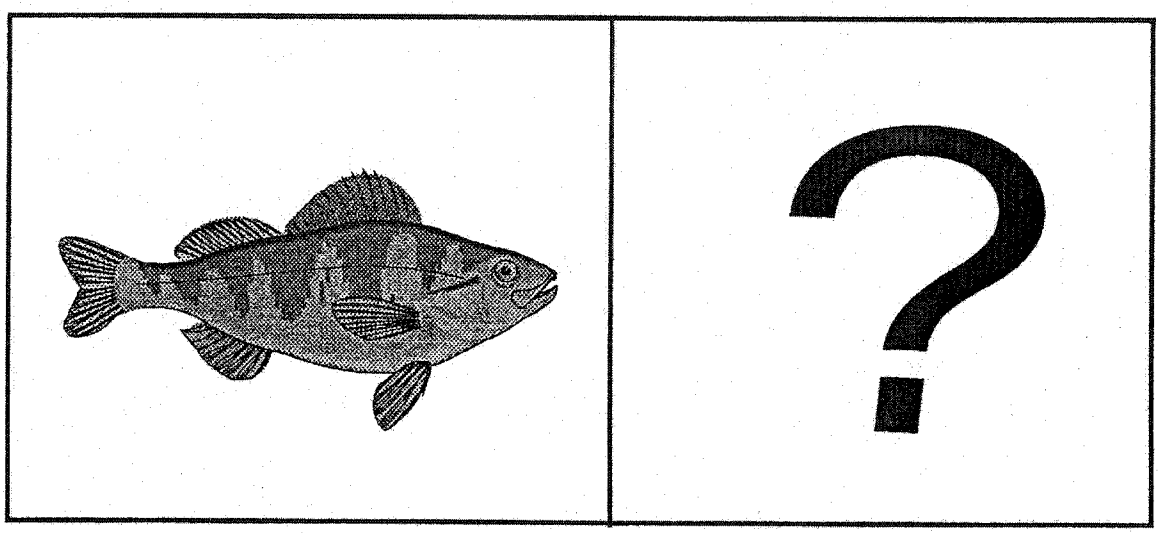


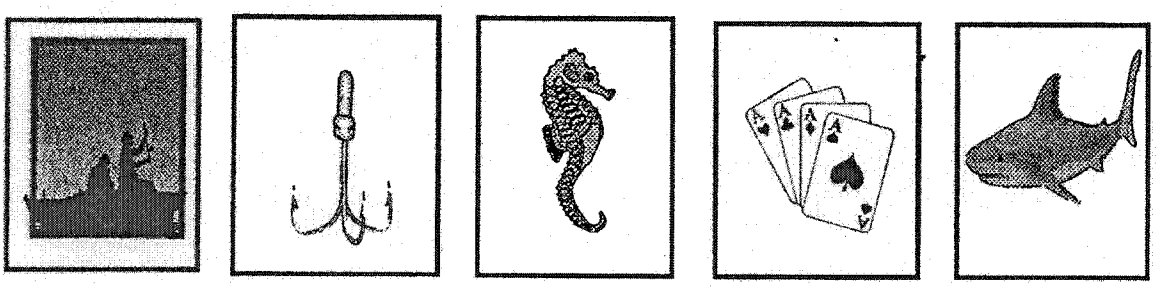
GAT (Version 1) Set A Question 16



GAT (Version 1) Set A Question 17

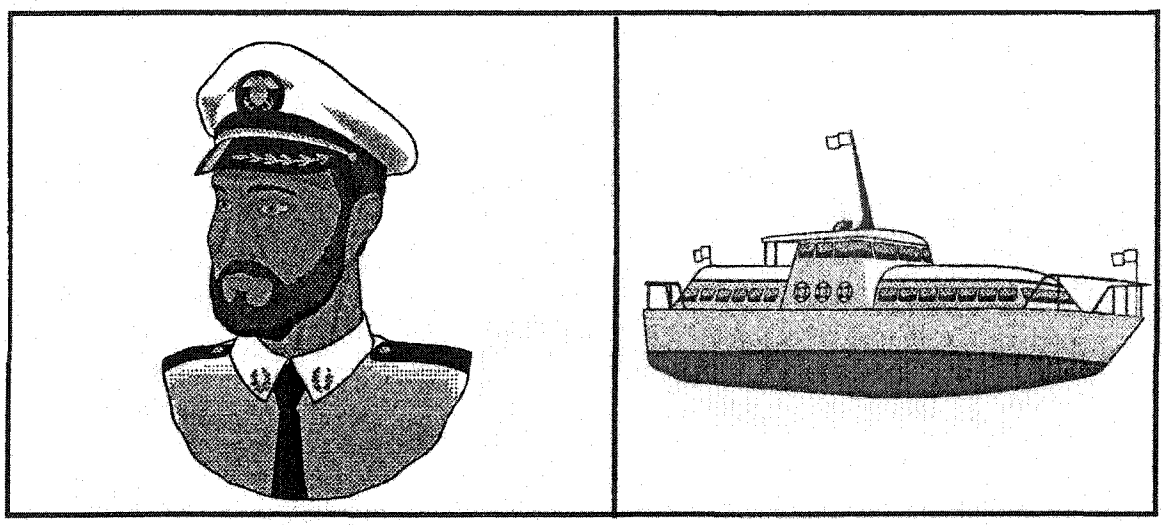


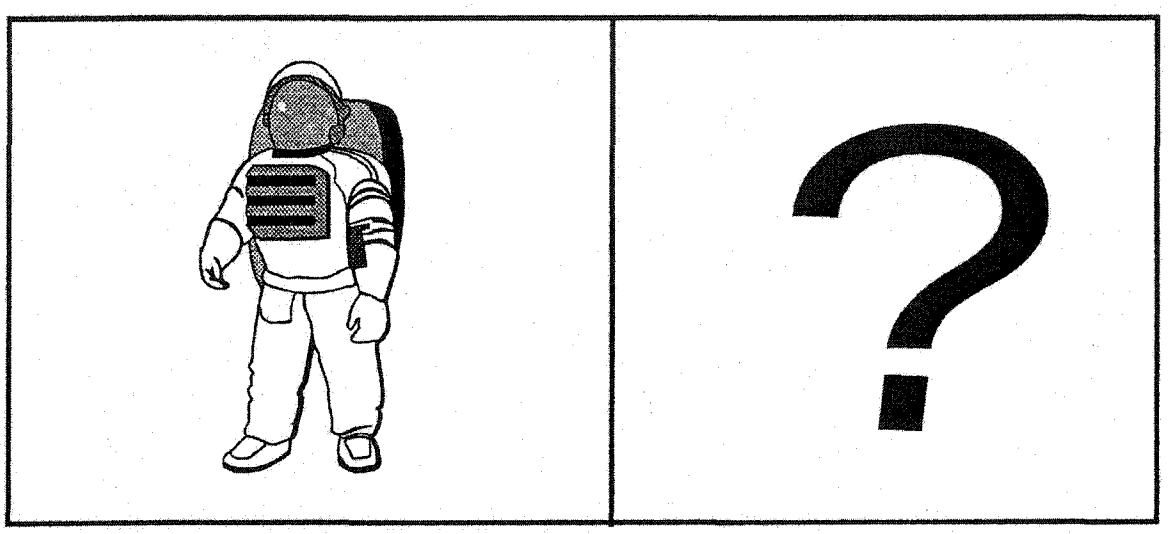


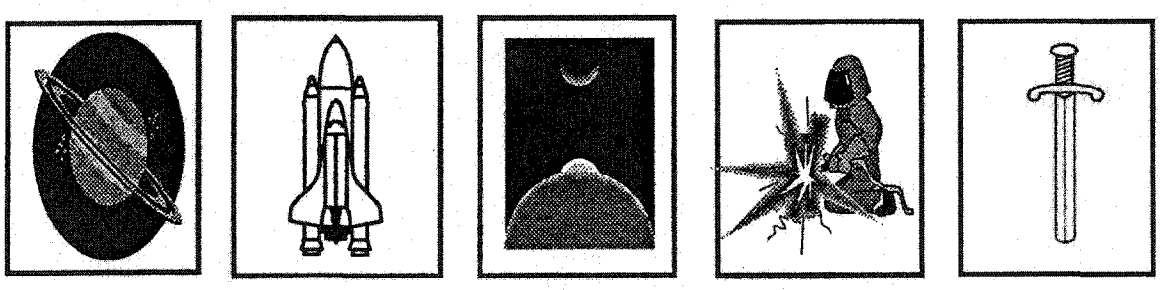


A B C D E

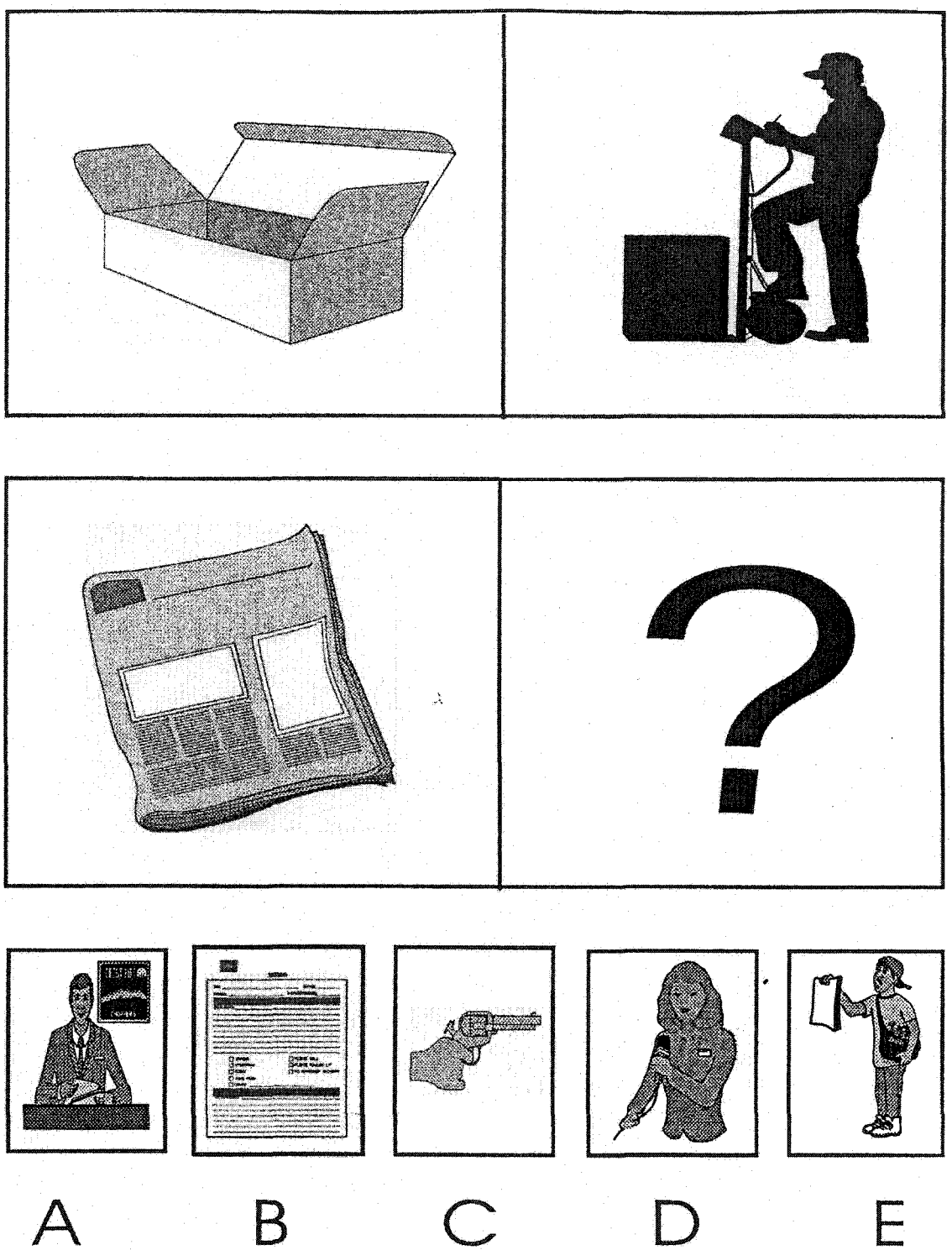
GAT (Version 1) Set A Question 18



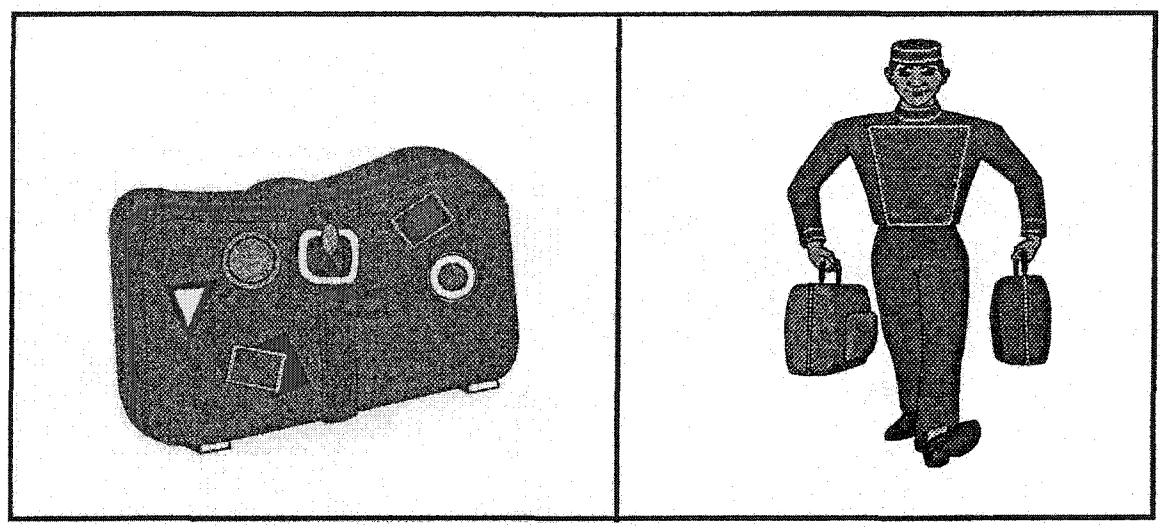


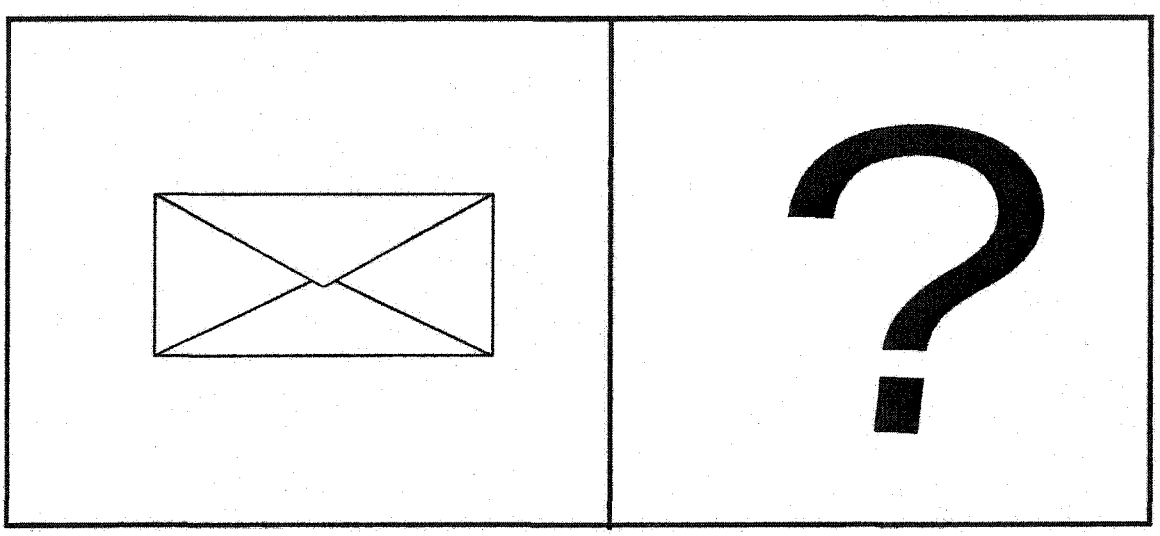


GAT (Version 1) Set A Question 19



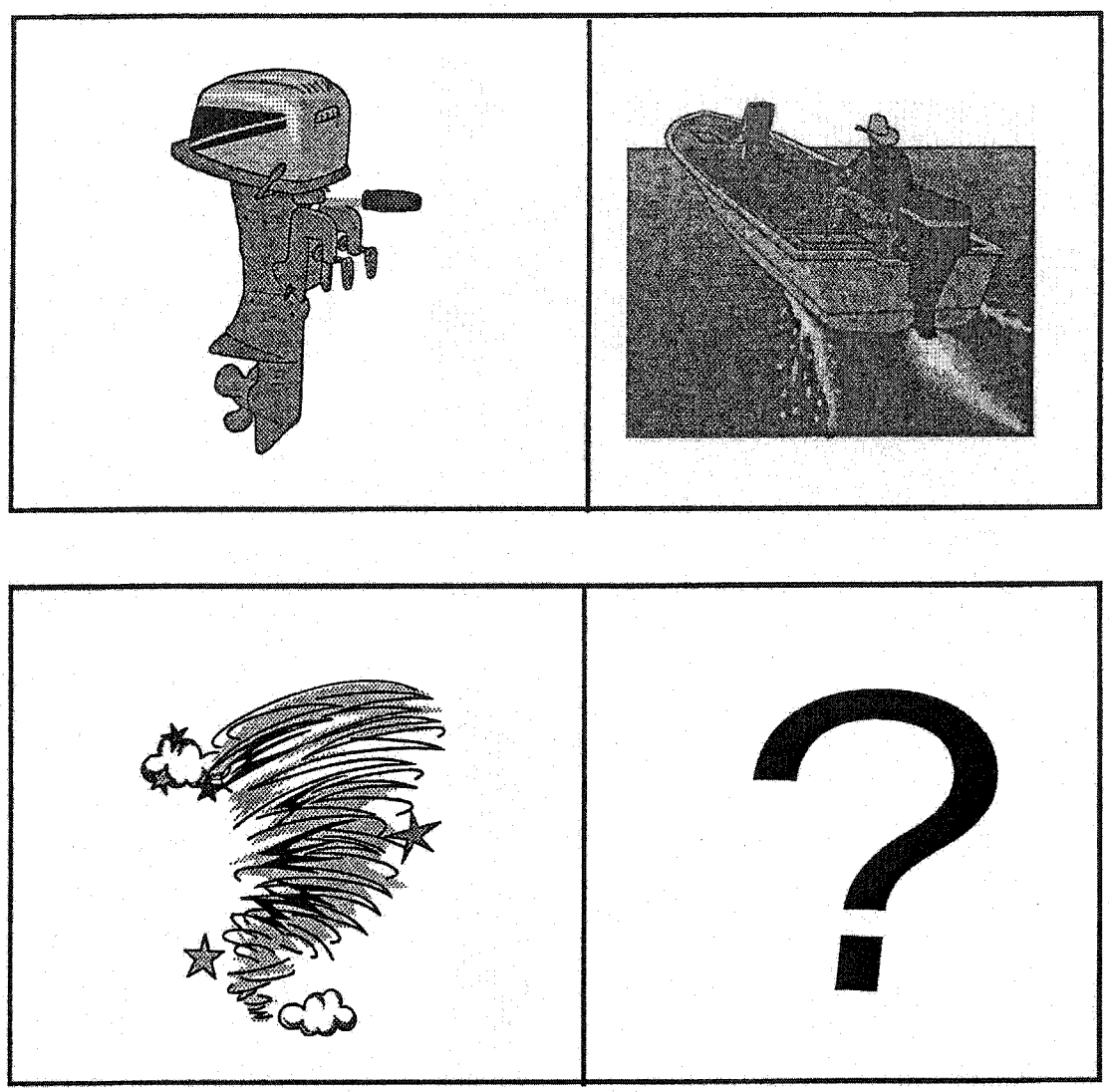
GAT (Version 1) Set A Question 20



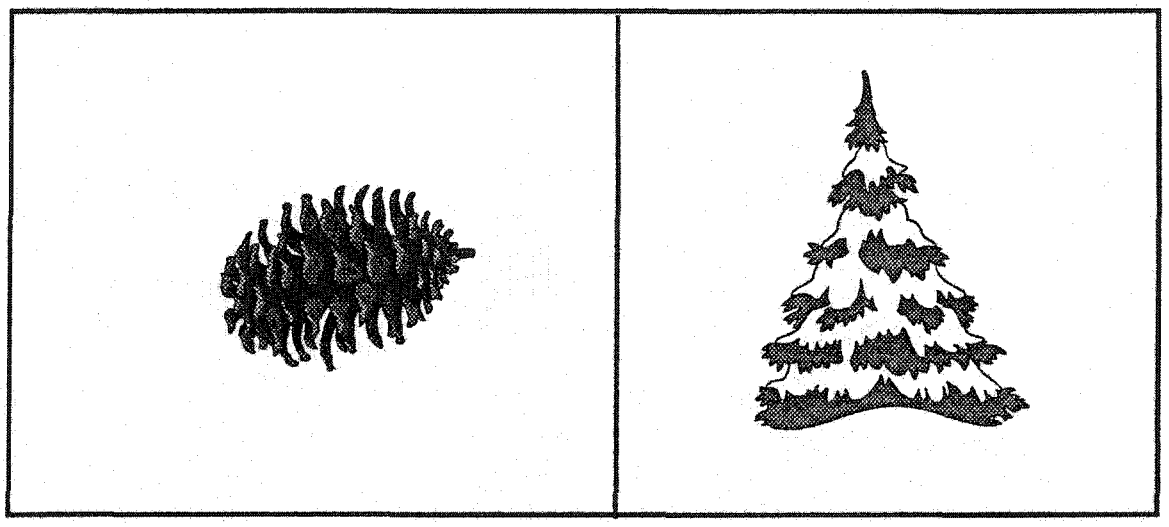


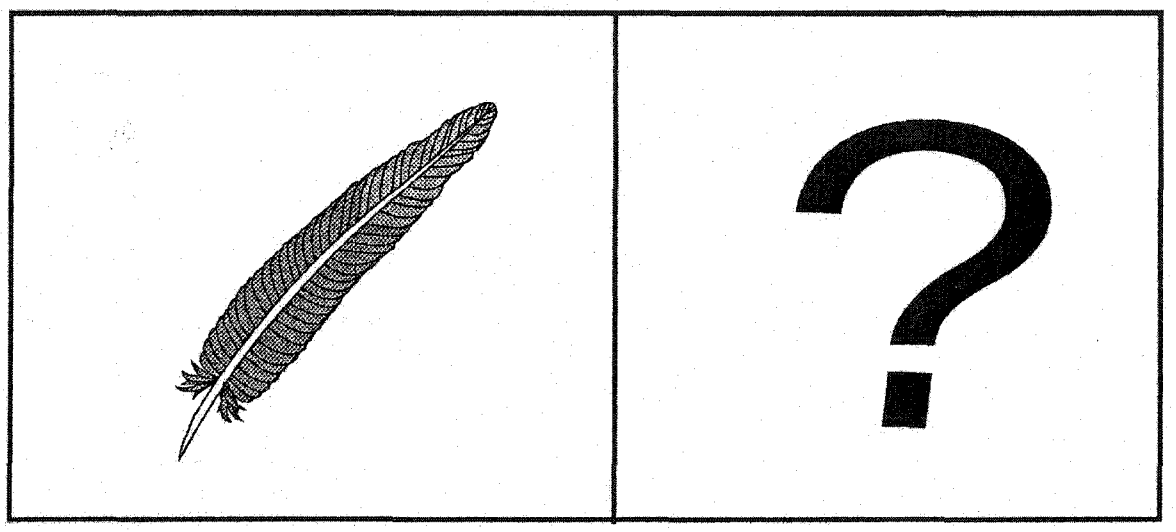
A B C D E

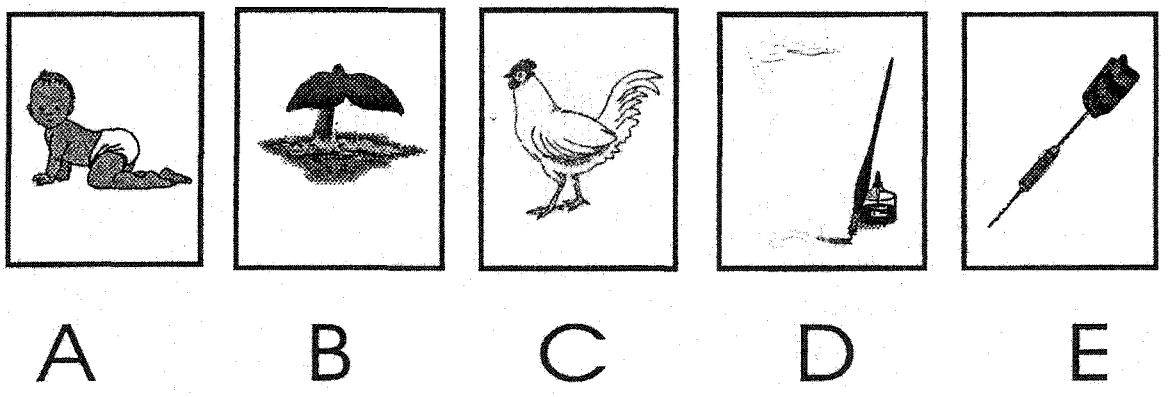
GAT (Version 1) Set A Question 21



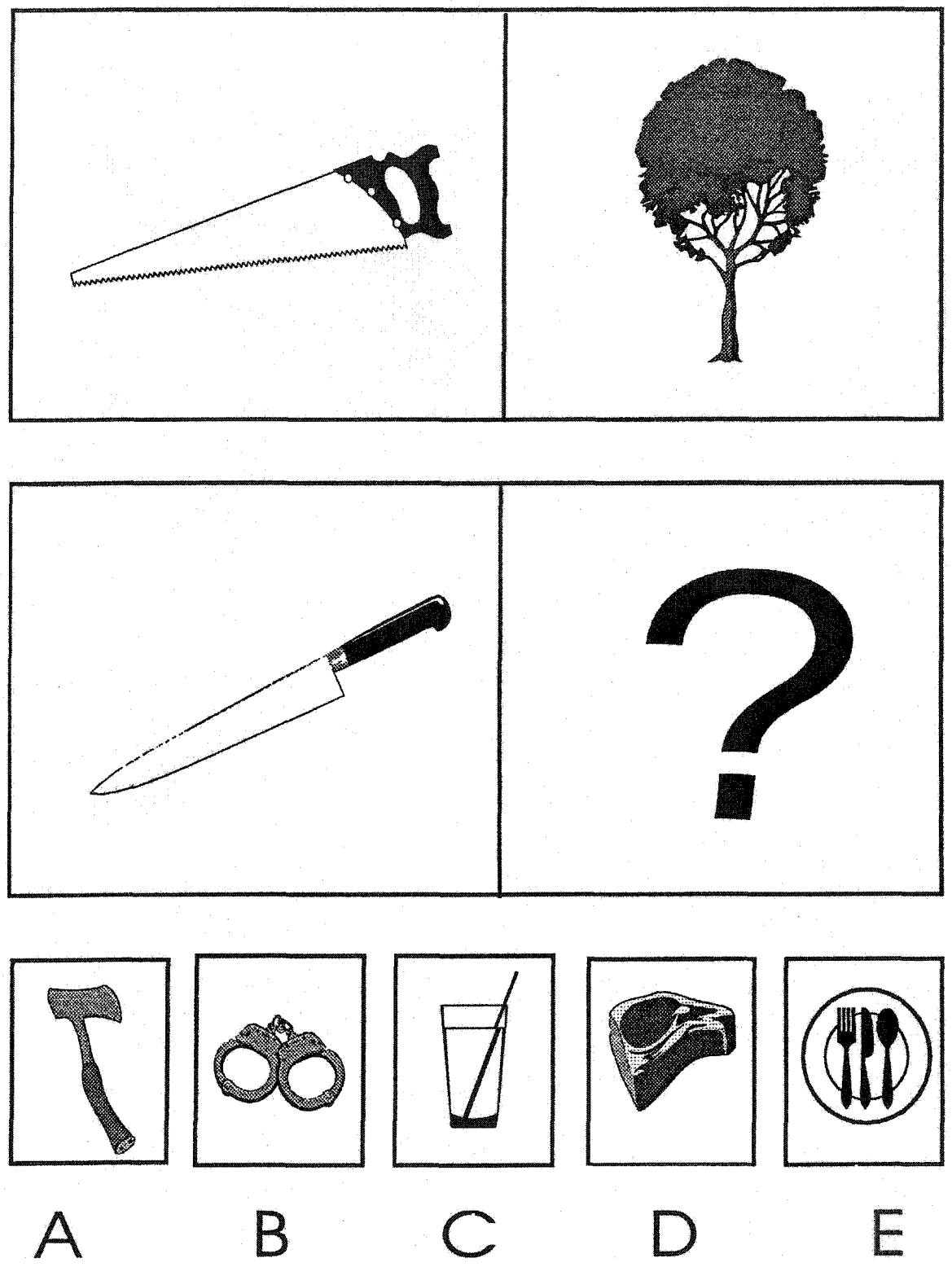
GAT (Version 1) Set A Question 22







GAT (Version 1) Set A Question 23



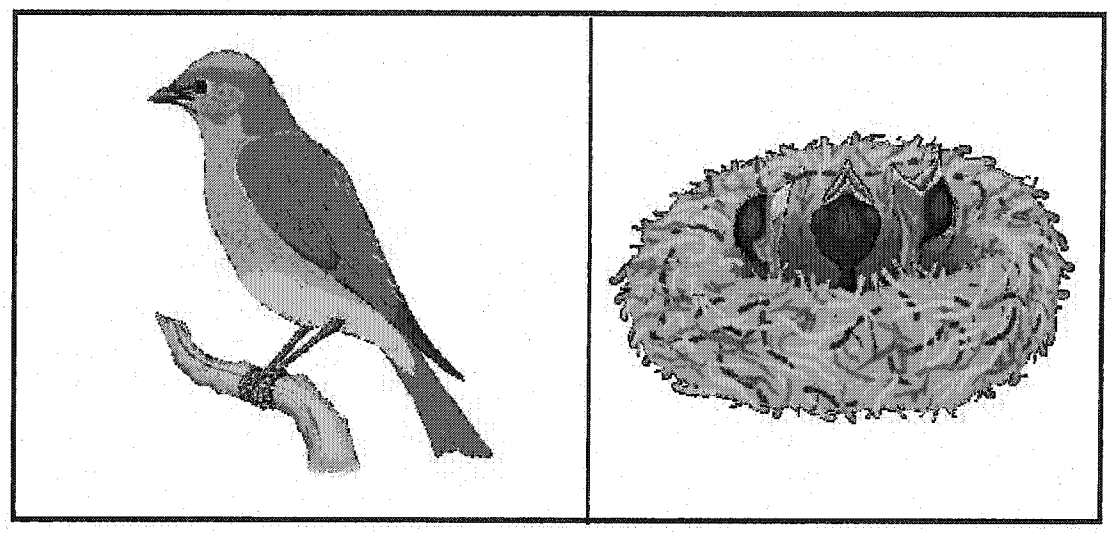
GAT (Version 1) Set A Question 24

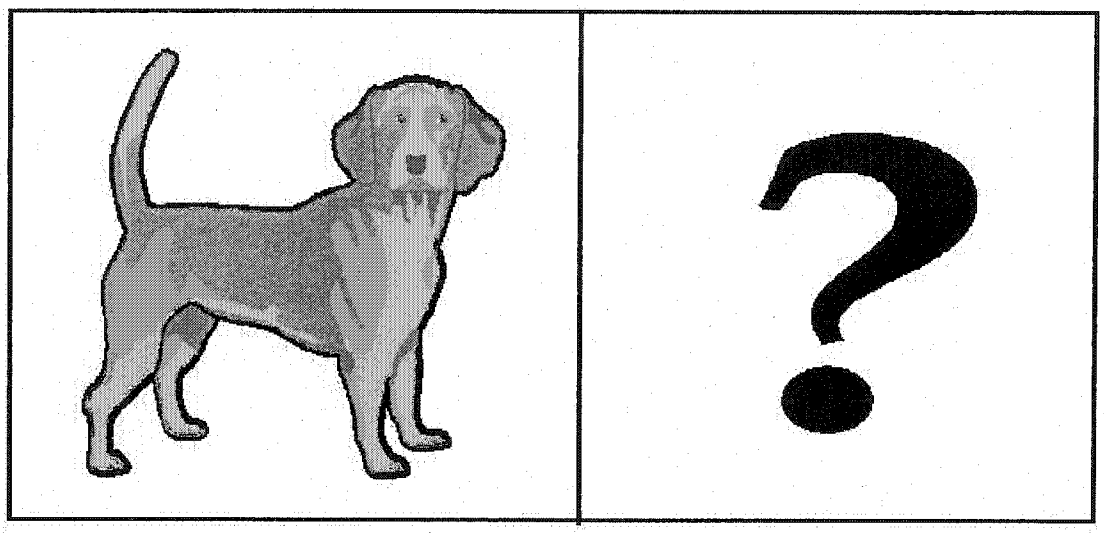
APPENDIX B

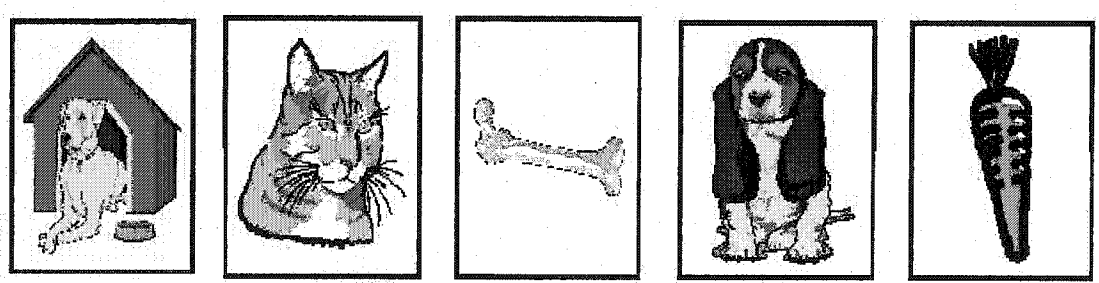
GORANSON ANALOGY TEST

(GAT: Version 2)

Test Developer:	Tamara Goranson, M.A. Ph.D. Candidate in Clinical Neuropsychology Department of Psychology University of Victoria Phone: (519) 721-7525 e-mail: tgoranso@uvic.ca
Ph.D. Supervisor:	Roger Graves, Ph.D., R. Psych. Department of Psychology University of Victoria Phone: (250) 721-7539 Fax: (250) 721-8929 e-mail: rgraves@uvic.ca
University Affiliation:	Department of Psychology University of Victoria P.O. Box 3050 Victoria, B.C., CANADA V8W 3P5







A B C D E

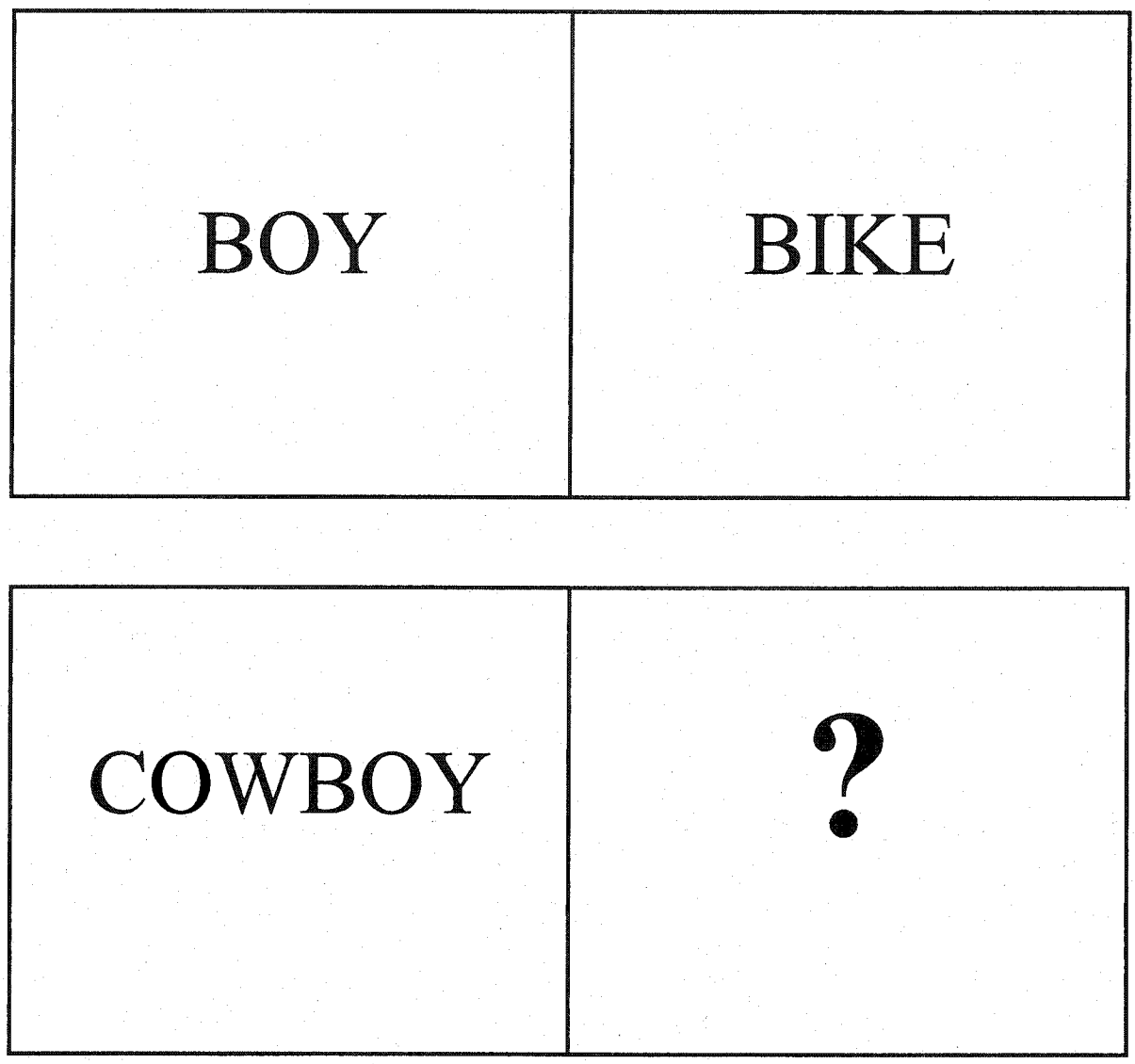
Practice Problem

SHEEP	WOOL SWEATER

COW	?

A	B	C	D	E
HORSE	CALF	GARBAGE CAN	LEATHER BOOT	MILK

PRACTICE PROBLEM



A	B	C	D	E
INDIAN	BOOT	BULL - FIGHTER	WHEEL - BARROW	HORSE

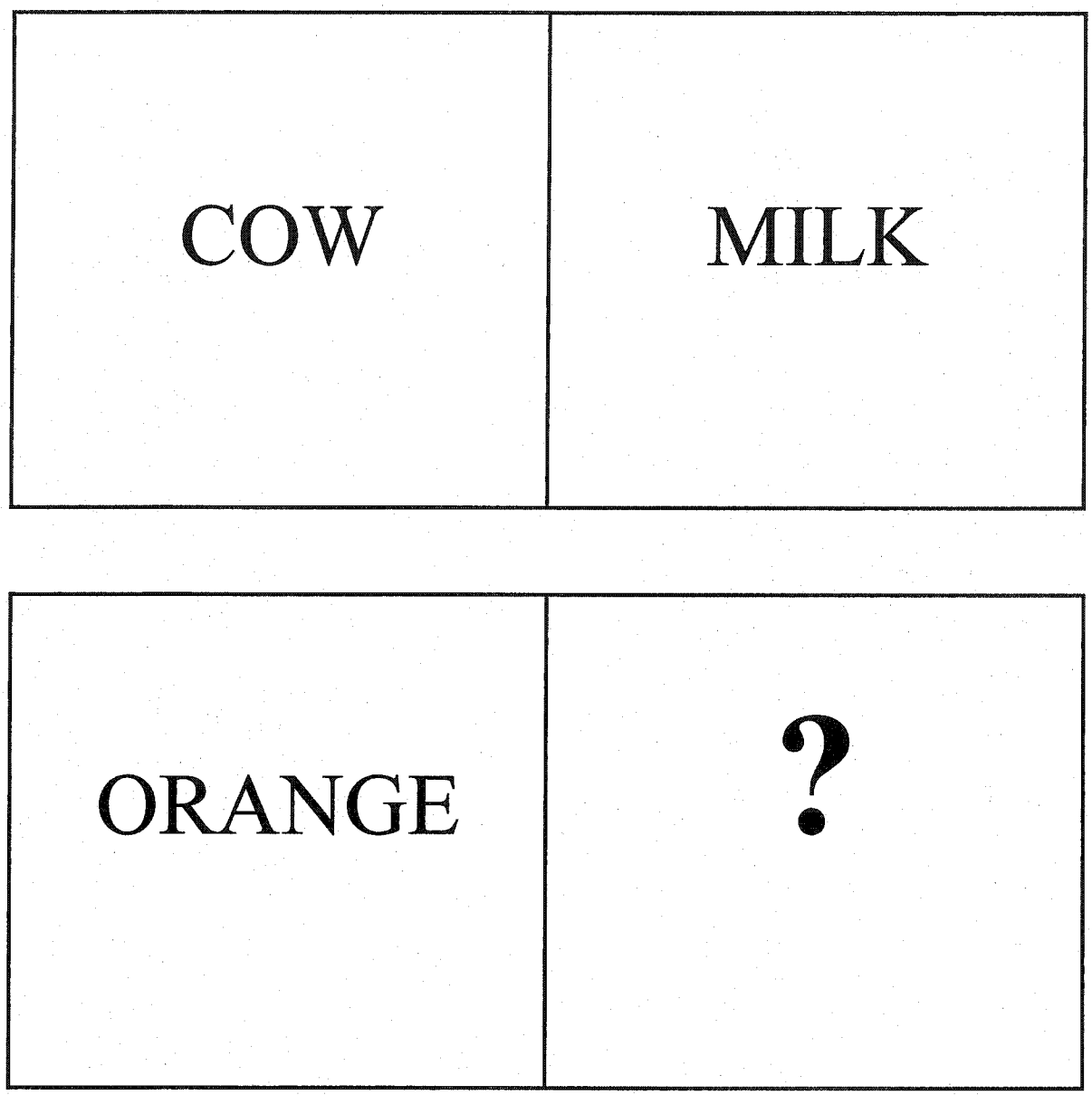
GAT (Version 2) Set A Question 1

BASEBALL	BASEBALL PLAYER

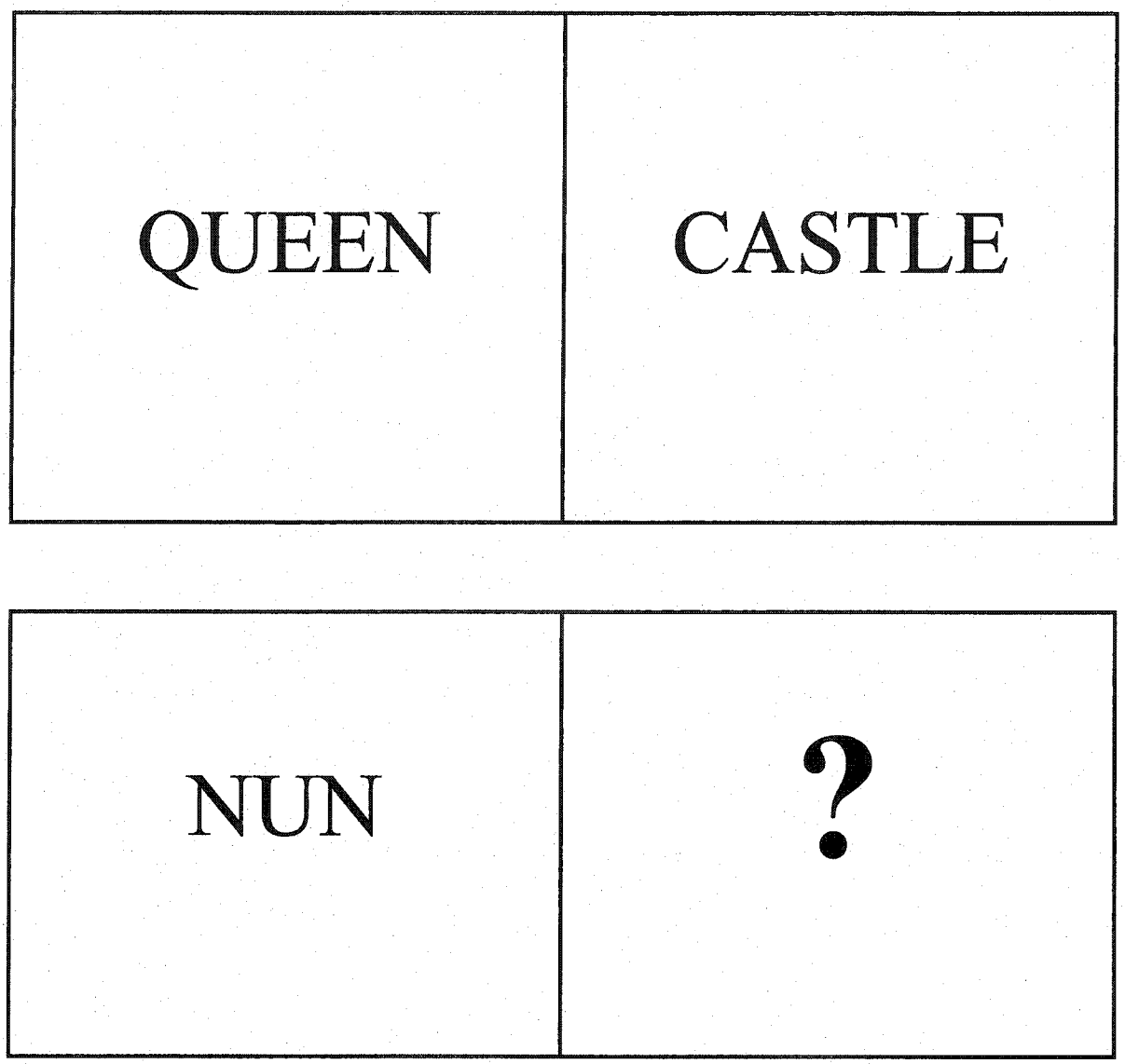
FOOTBALL	?

A	B	C	D	E
TELEPHONE	FOOTBALL PLAYER	FOOTBALL HELMET	BASKETBALL	LEMON

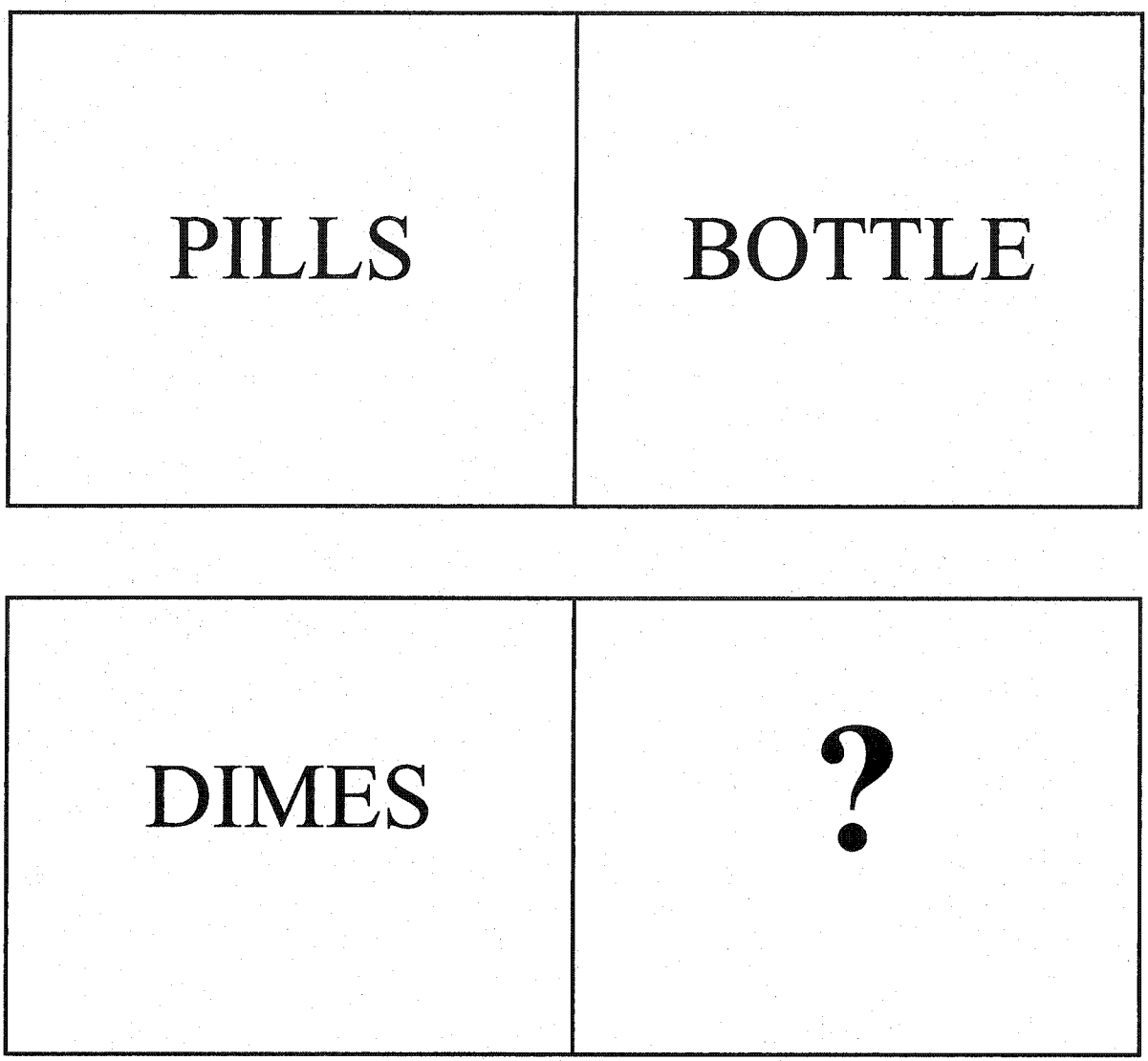
GAT (Version 2) Set A Question 2



GAT (Version 2) Set A Question 3

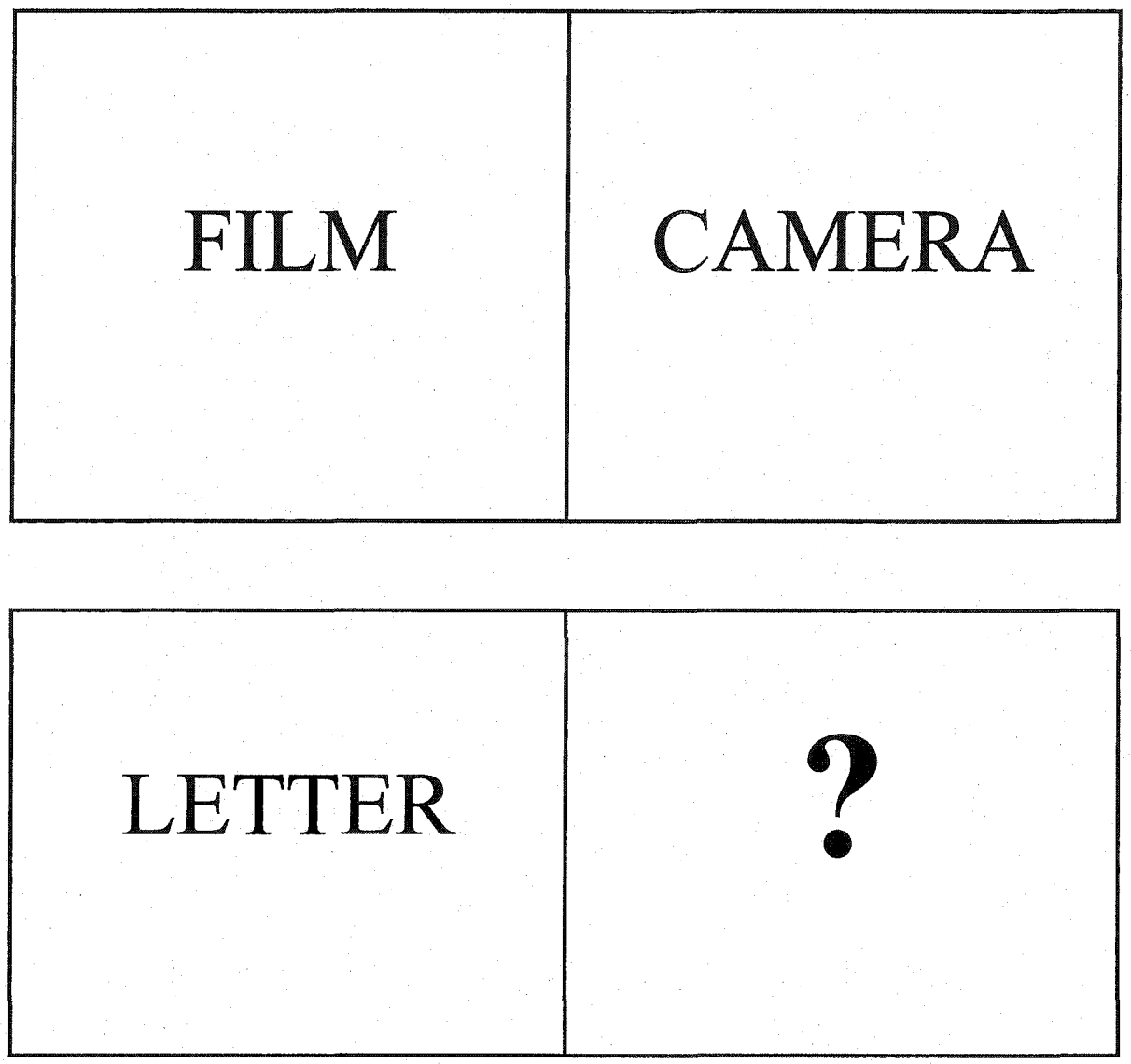


GAT (Version 2) Set A Question 4

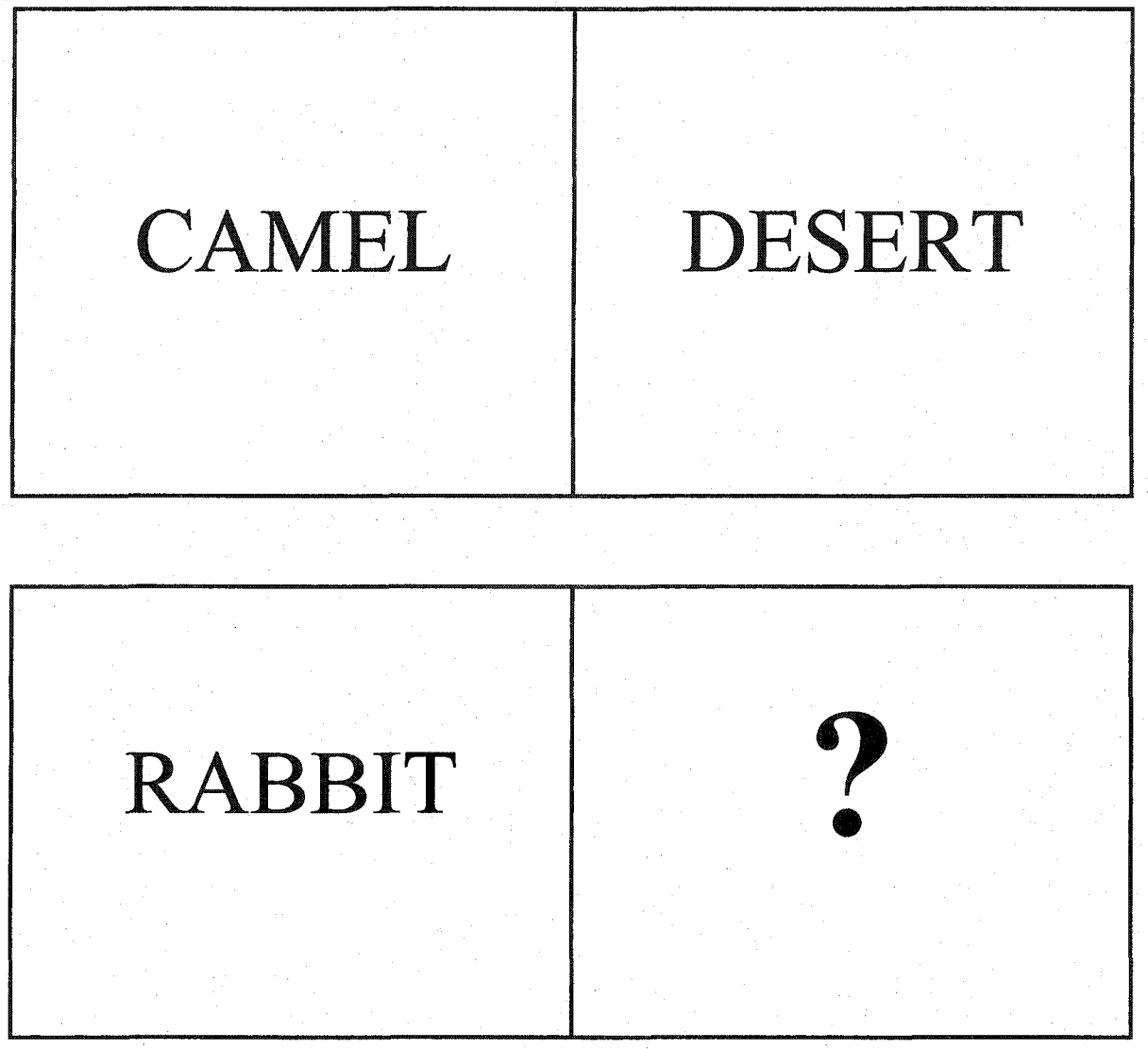


CREDIT CARD	PENNIES	BAND-AID	PIGGY-BANK	MONEY
A	B	C	D	E

GAT (Version 2) Set A Question 5

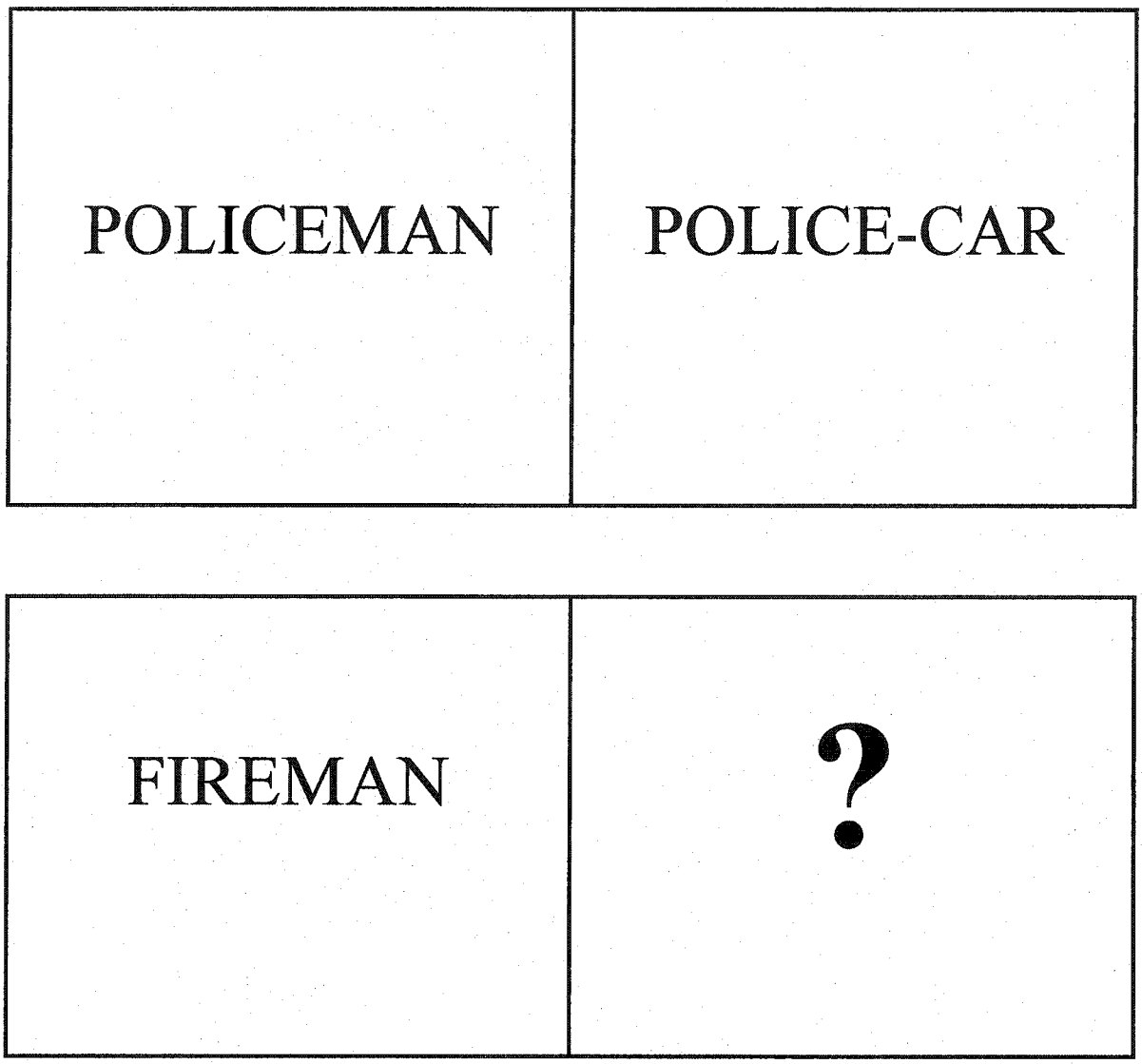


GAT (Version 2) Set A Question 6



PAIL	FOREST	CARROT	DEER	BUNNY
A	B	C	D	E

GAT (Version 2) Set A Question 7



A	B	C	D	E
FIRE EXTINGUISHER	CONSTRUCTION WORKER	SUITCASE	FIRE TRUCK	FIRE

GAT (Version 2) Set A Question 8

TRAIN	RAILWAY CROSSING
CAR	?

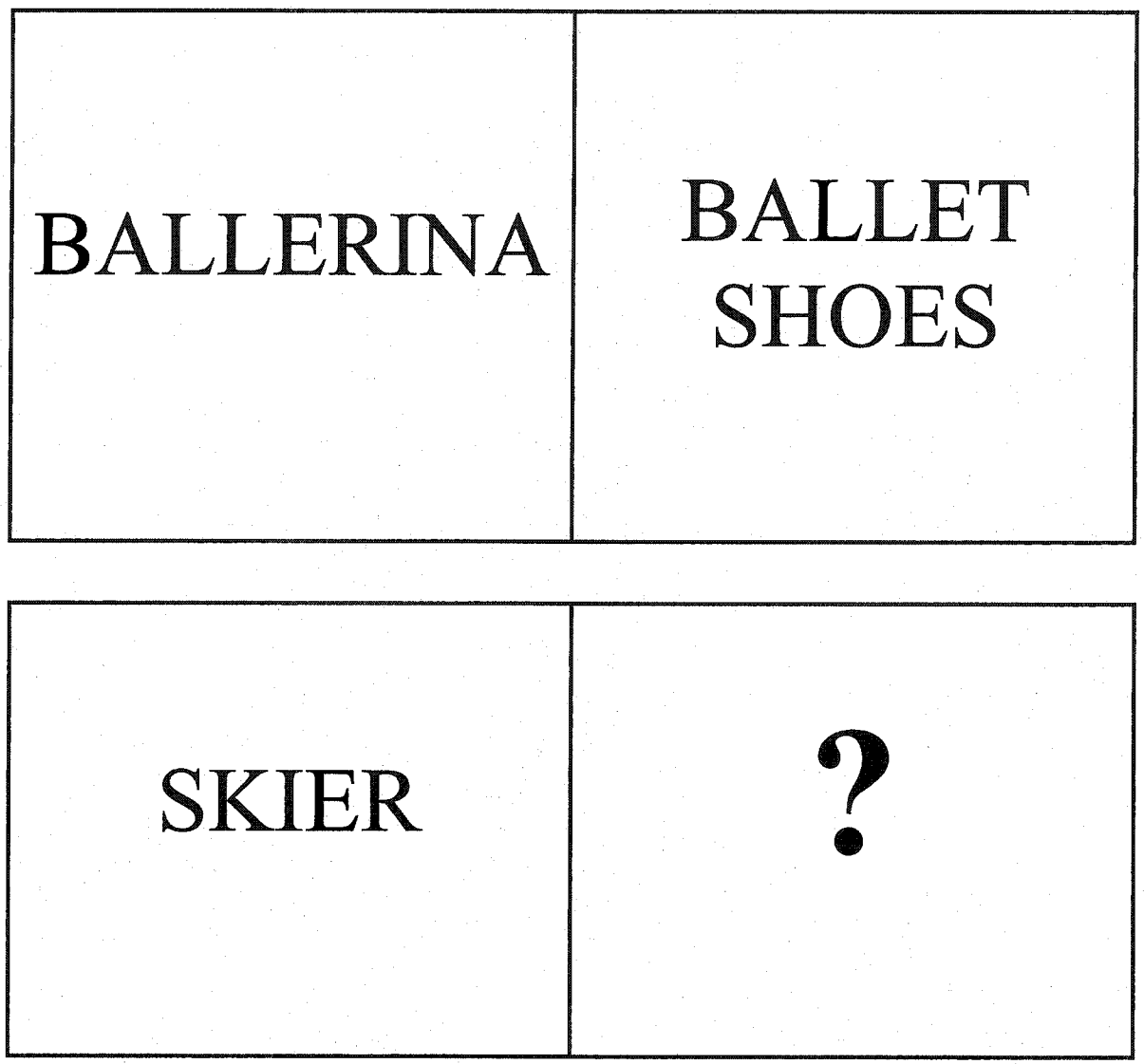
A	B	C	D	E
WHEEL - CHAIR	TRAFFIC LIGHTS	TRUCK	AIRPLANE	JEEP

GAT (Version 2) Set A Question 9

NURSE	NEEDLE
ARTIST	?

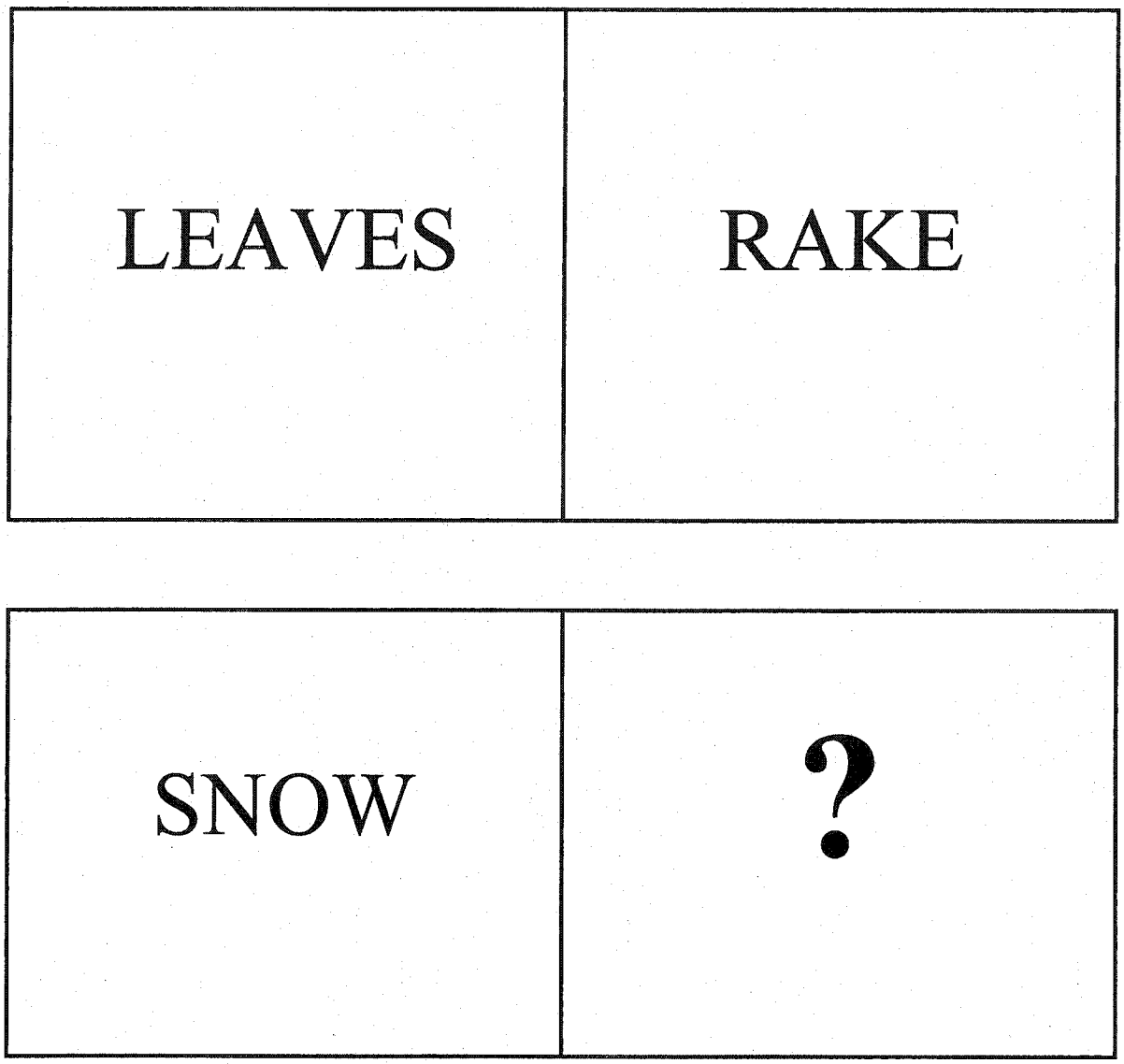
A	B	C	D	E
PAINT - BRUSH	PAINT	PENCIL	PAINTER	KEY

GAT (Version 2) Set A Question 10

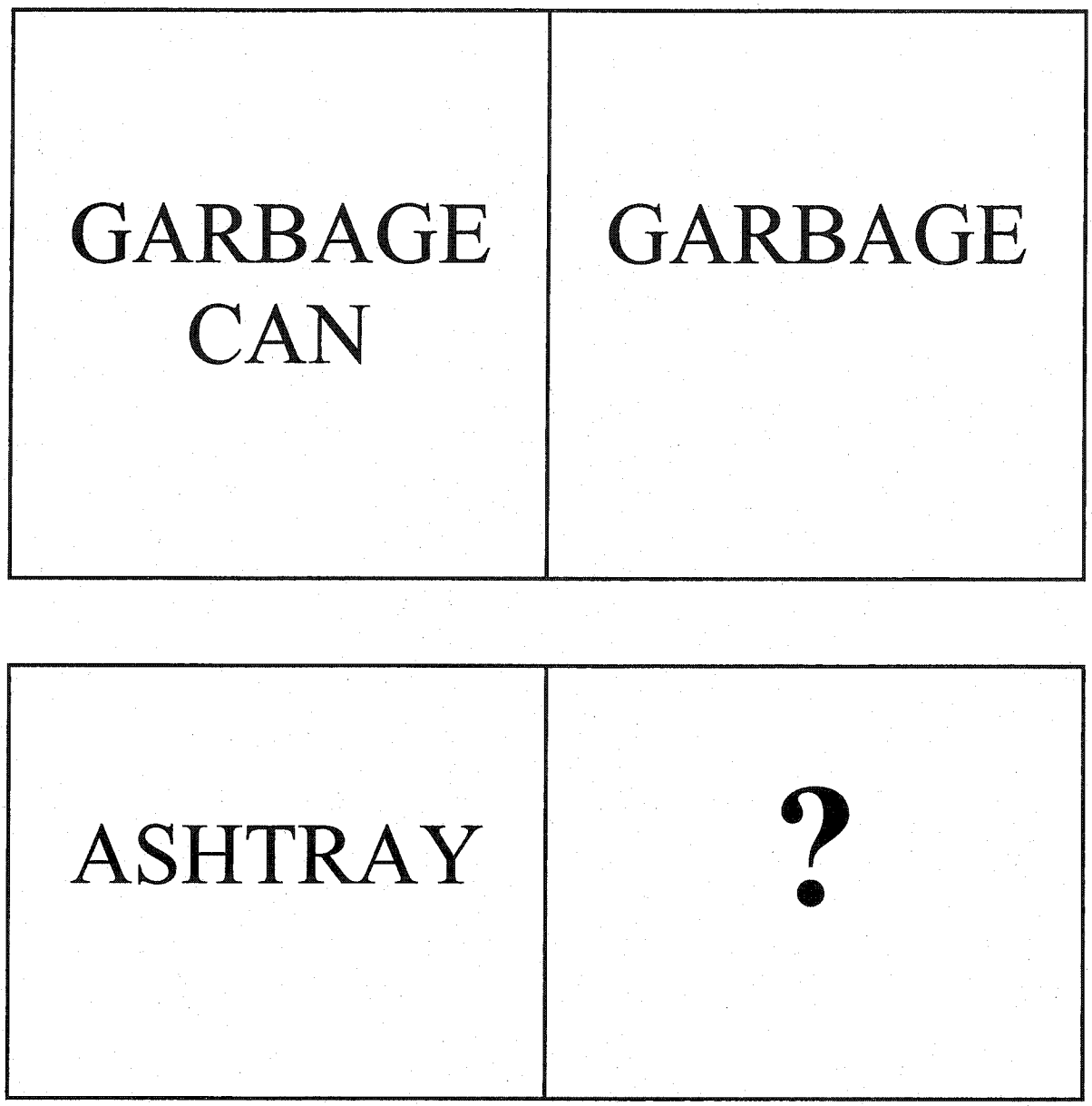


A	B	C	D	E
MOUNTAIN	SKATER	WATER-SKIER	PAPER-CLIP	SKI BOOT

GAT (Version 2) Set A Question 11

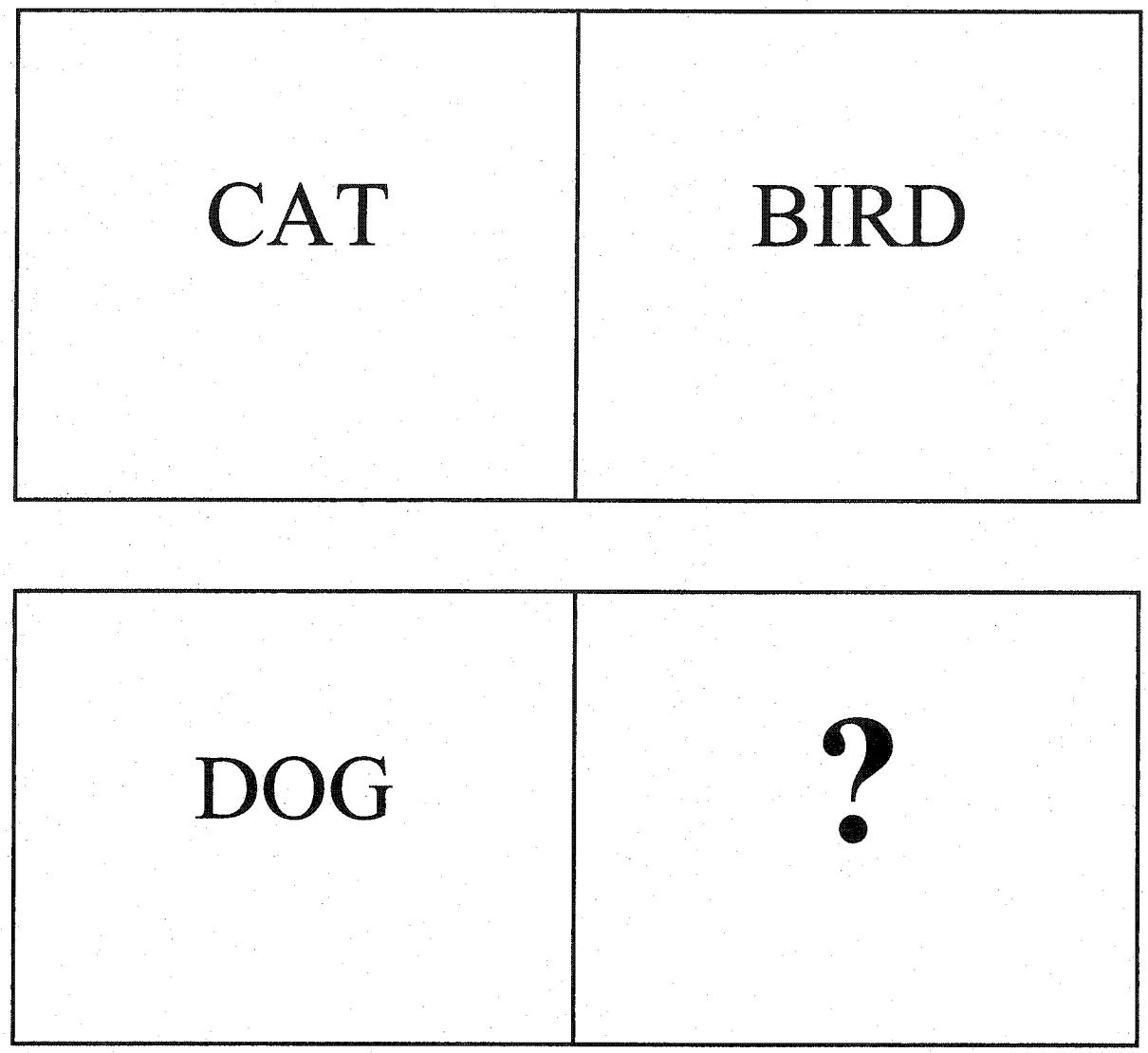


GAT (Version 2) Set A Question 12

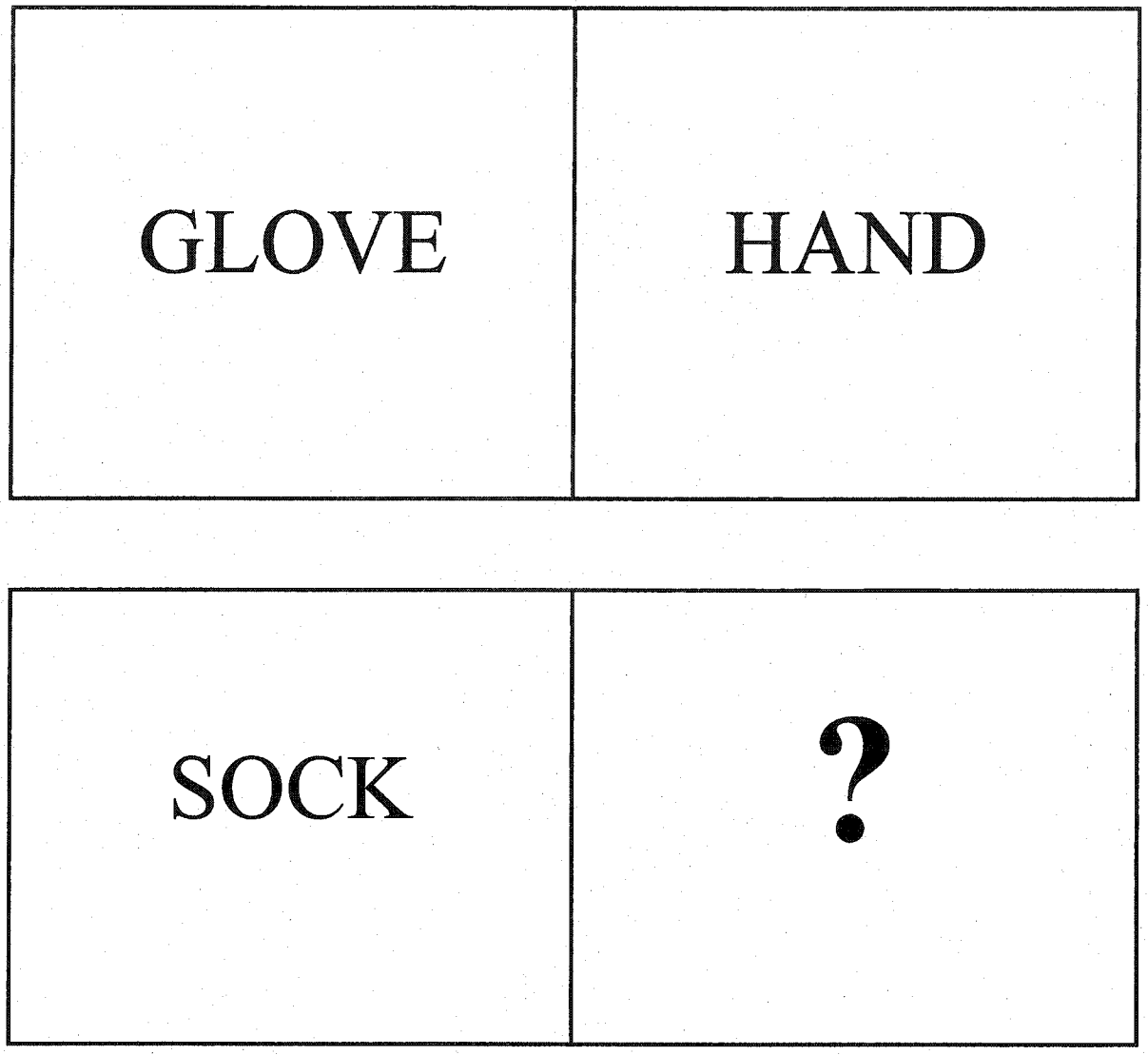


A	B	C	D	E
LIGHTER	BUTT	CIGARETTES	BOWL	BROCCOLI

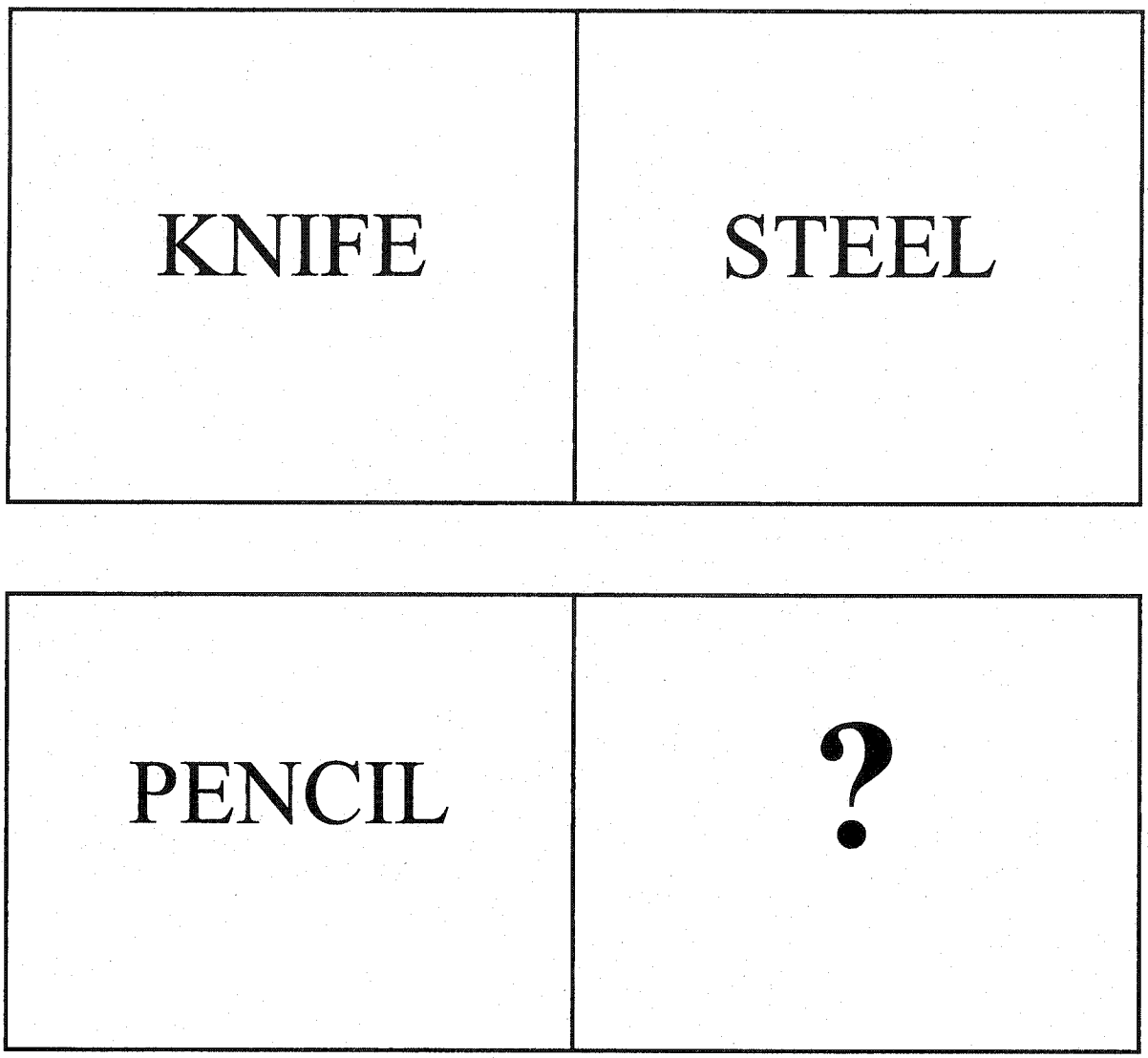
GAT (Version 2) Set A Question 13



GAT (Version 2) Set A Question 14

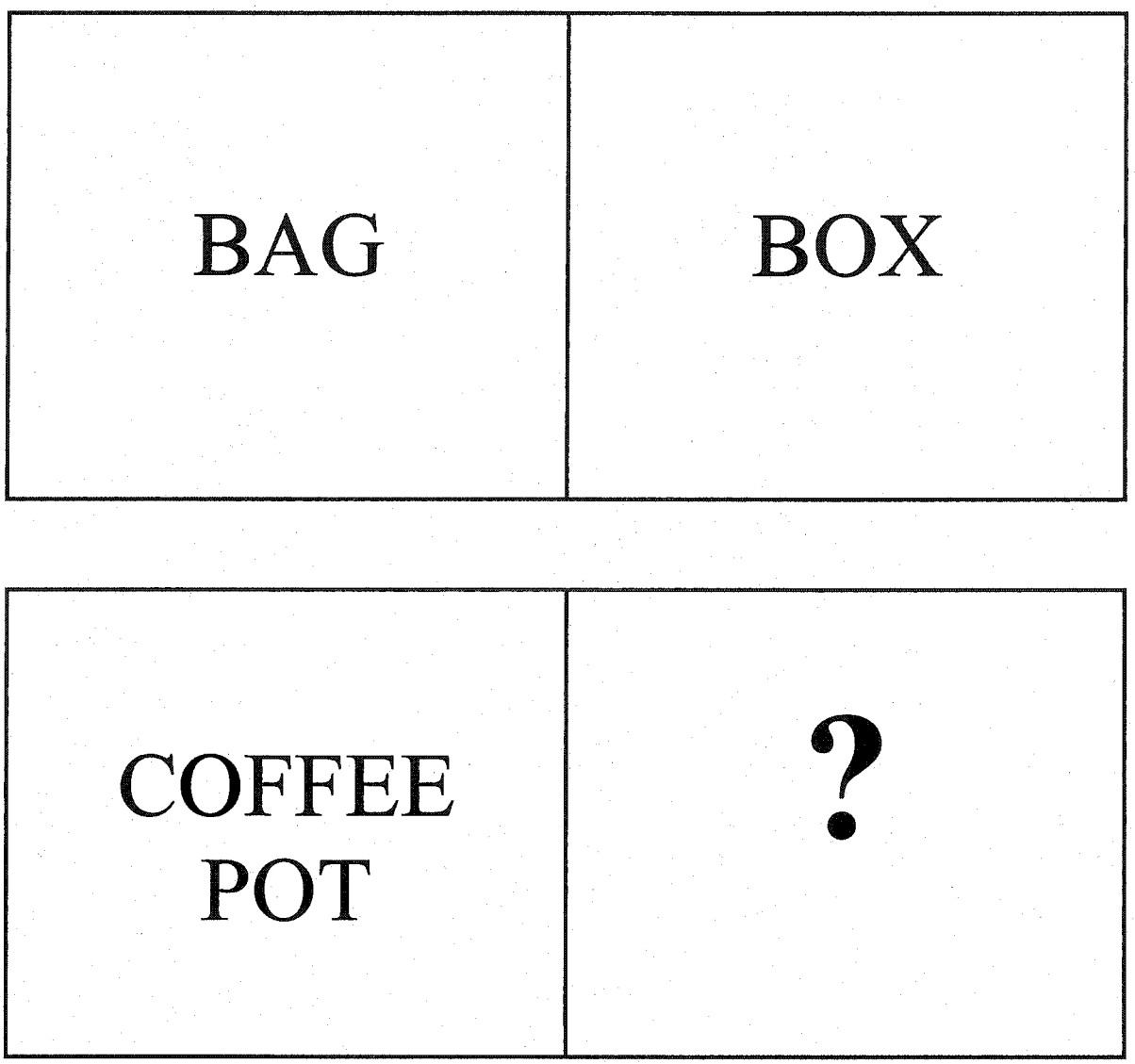


GAT (Version 2) Set A Question 15



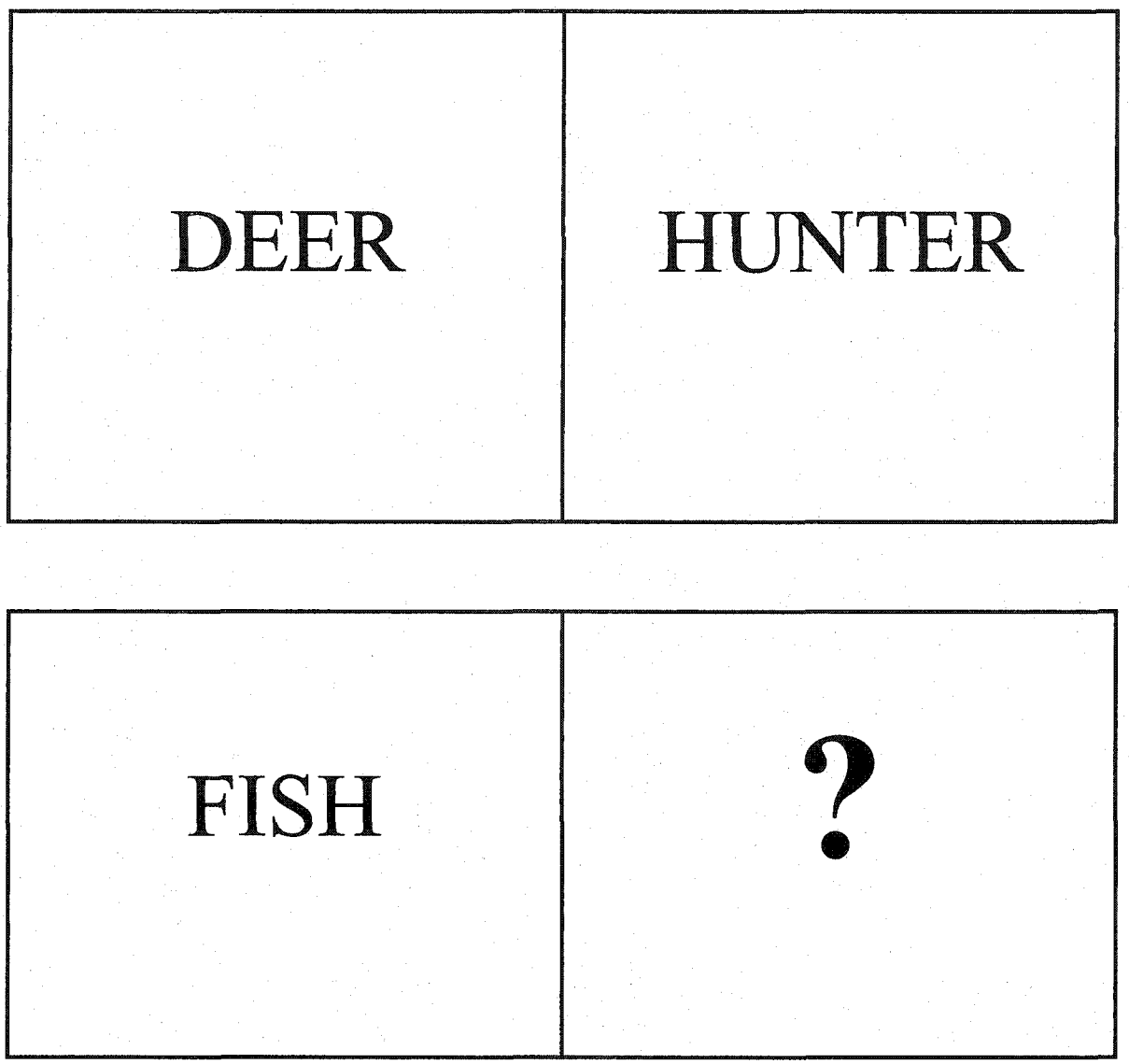
A	B	C	D	E
PHONE	STAPLER	SHARPENER	PEN	CRAYON

GAT (Version 2) Set A Question 16



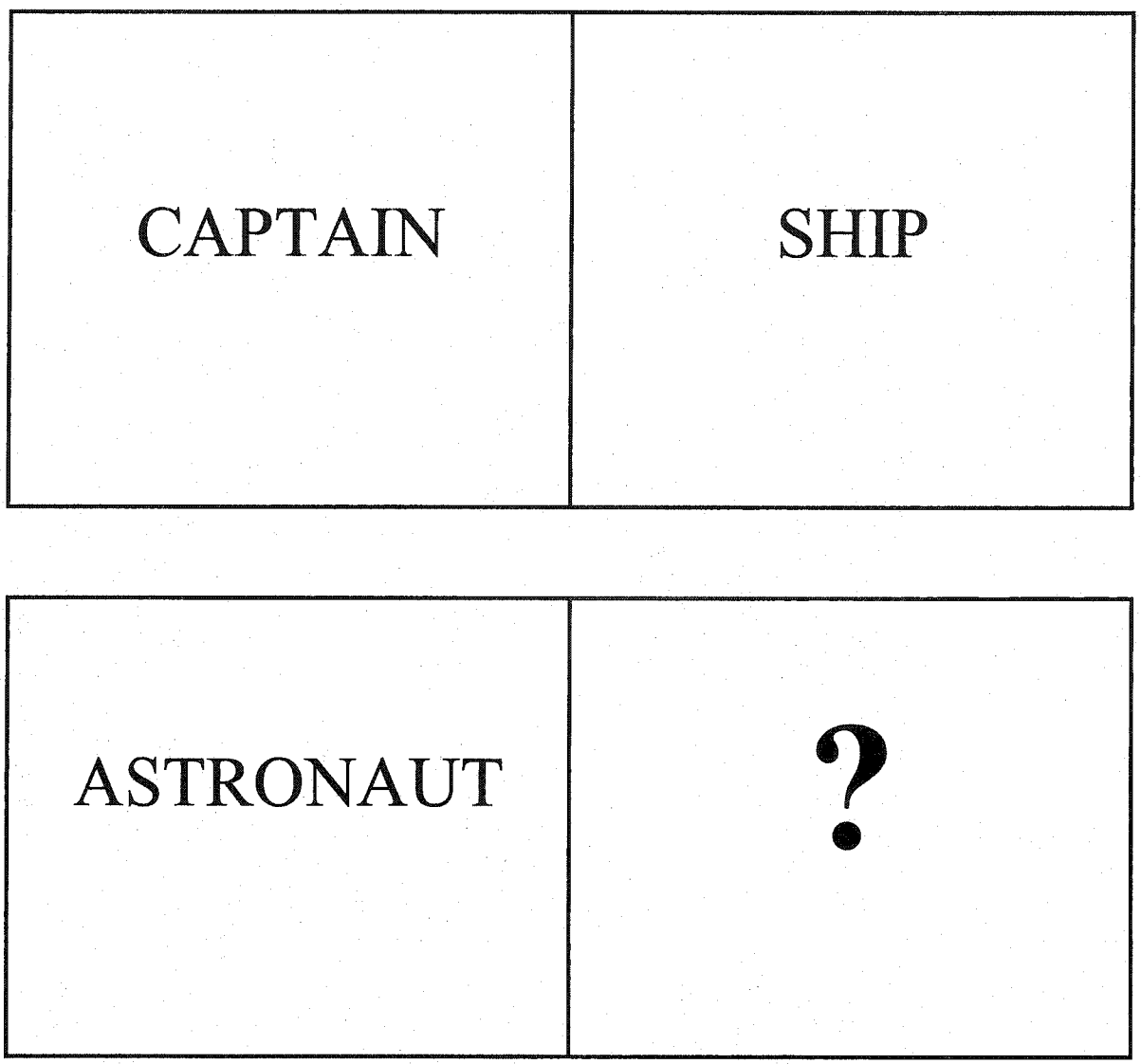
A	B	C	D	E
FIRE EXTINGUISHER	RING	UTENSILS	TEAPOT	COFFEE

GAT (Version 2) Set A Question 17



A	B	C	D	E
FISHERMAN	HOOK	SEAHORSE	CARDS	SHARK

GAT (Version 2) Set A Question 18

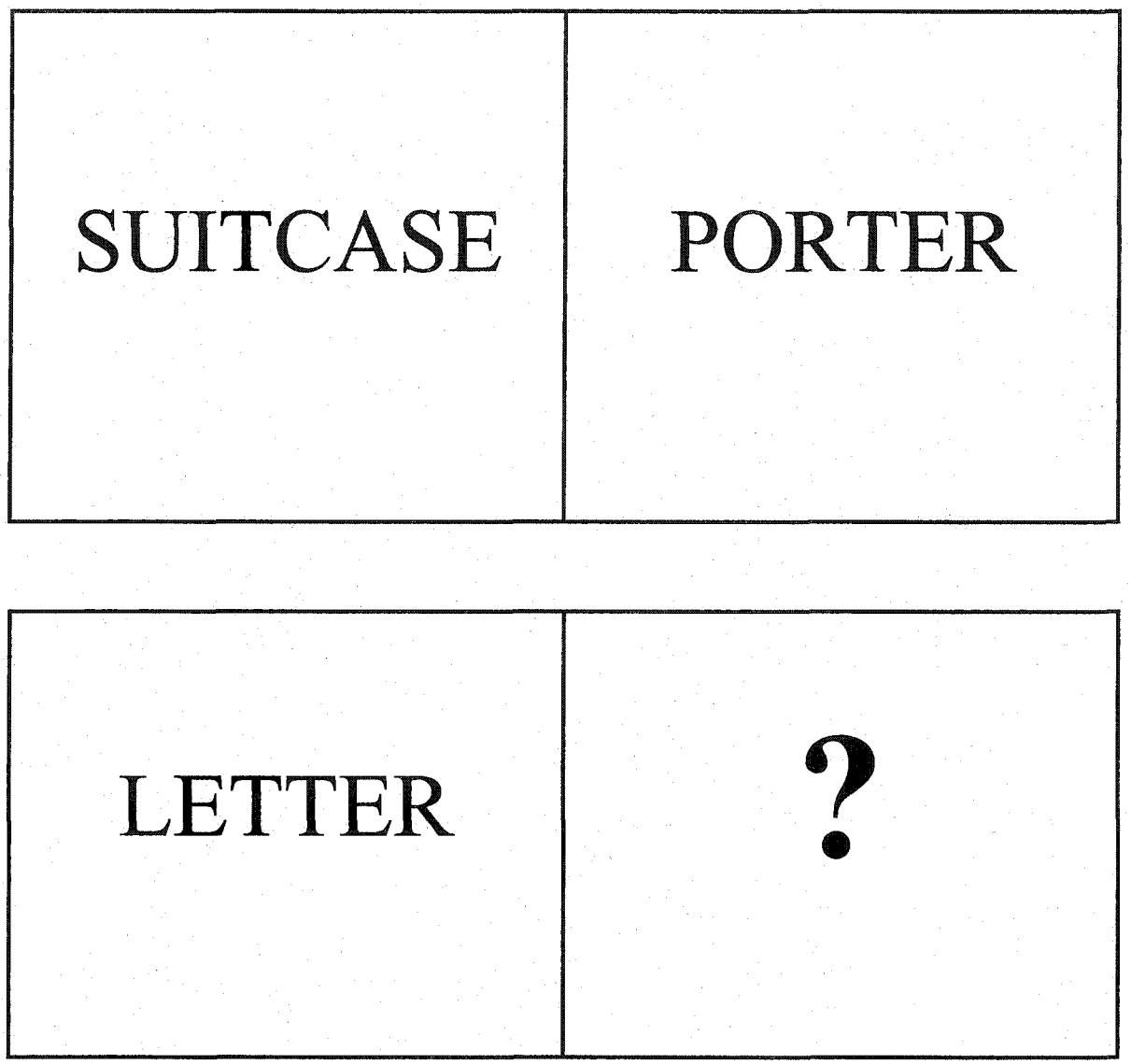


GAT (Version 2) Set A Question 19

BOX	DELIVERY MAN
NEWSPAPER	?

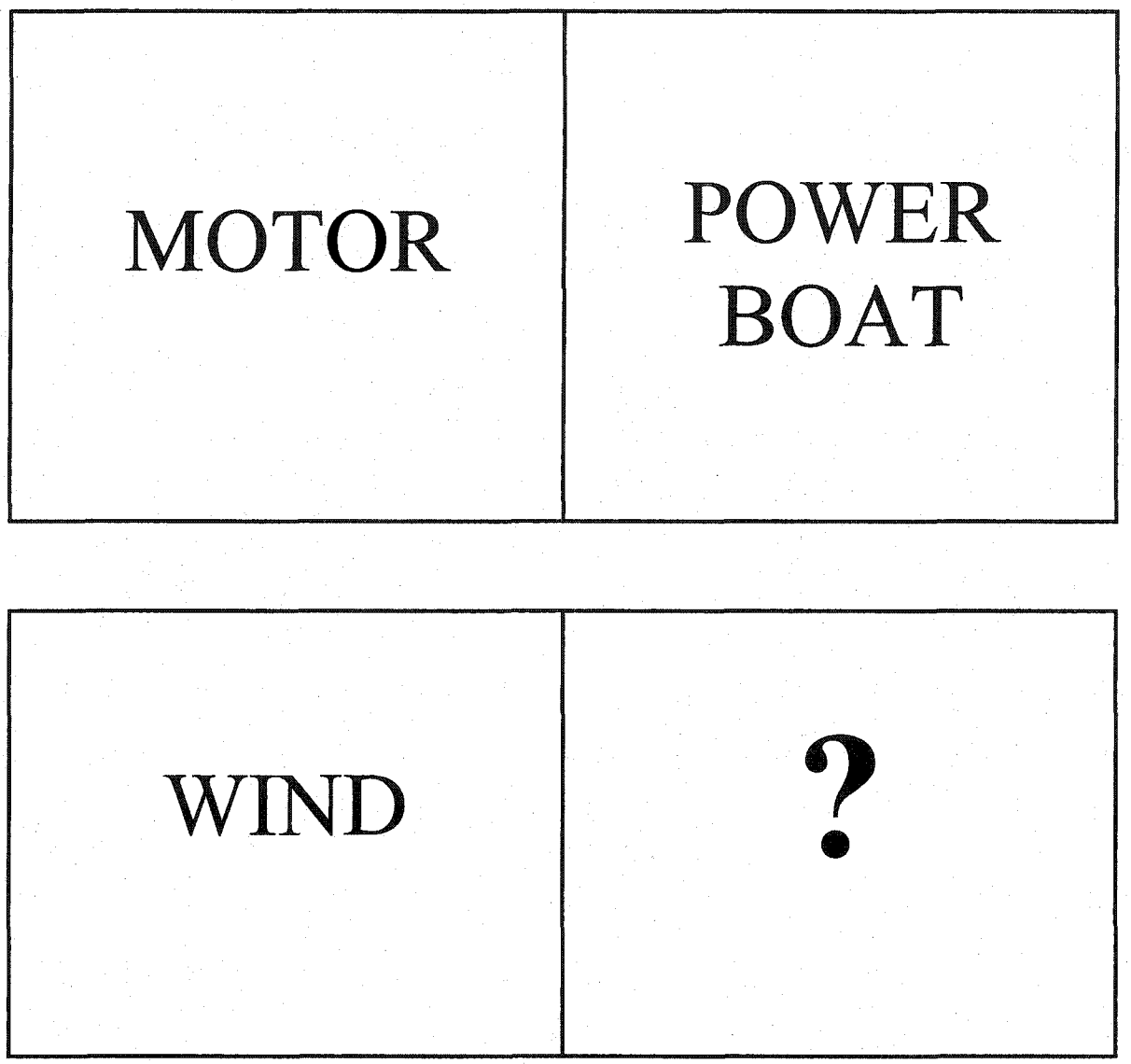
A	B	C	D	E
NEWS	REPORT	GUN	REPORTER	PAPERBOY

GAT (Version 2) Set A Question 20



A	B	C	D	E
CLOCK	STAMP	MAILMAN	MAILBOX	POSTCARD

GAT (Version 2) Set A Question 21



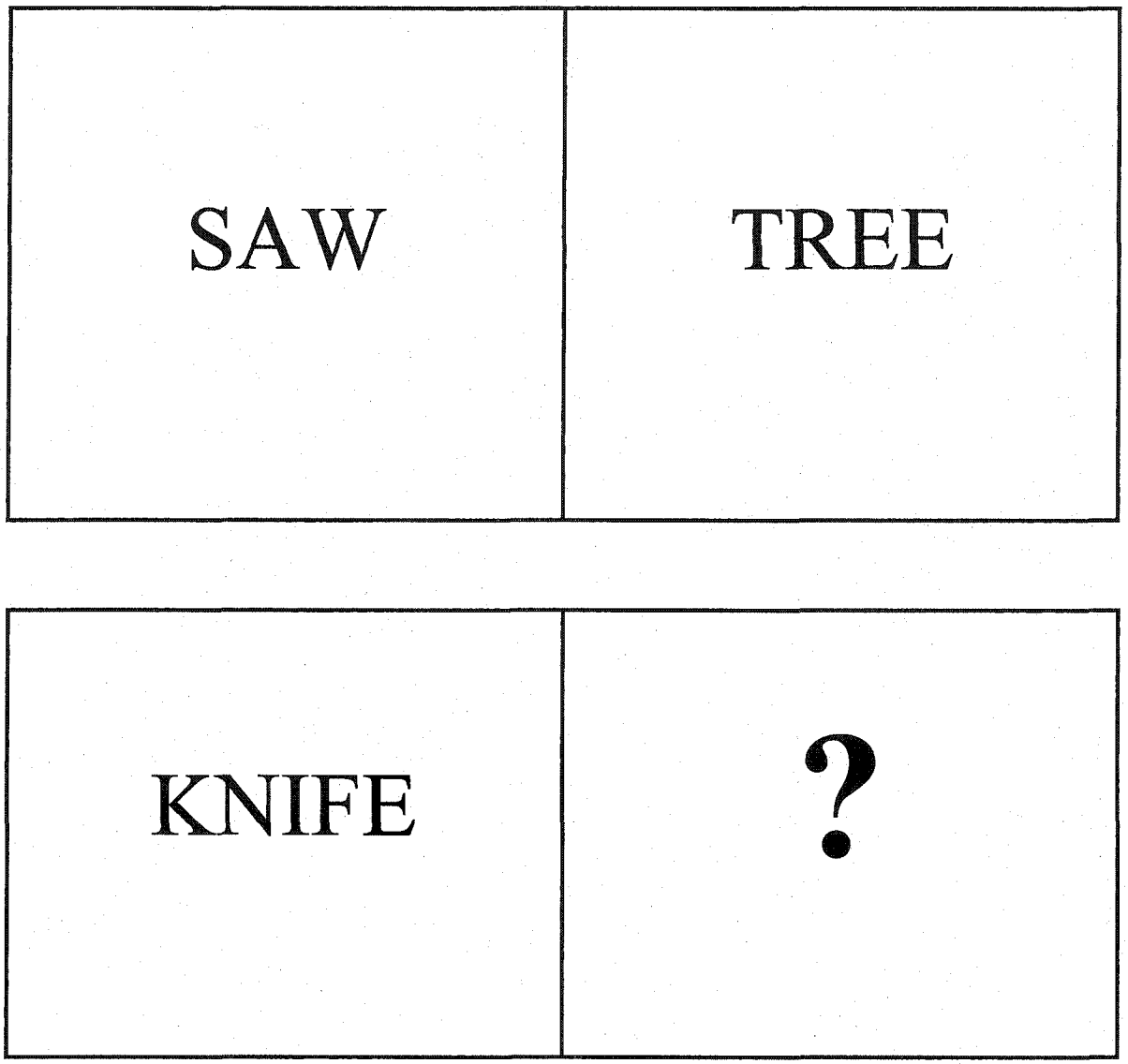
A	B	C	D	E
RAIN	SAILBOAT	WINDMILL	SMOKE	BALLOONS

GAT (Version 2) Set A Question 22

PINECONE	PINE TREE
FEATHER	?

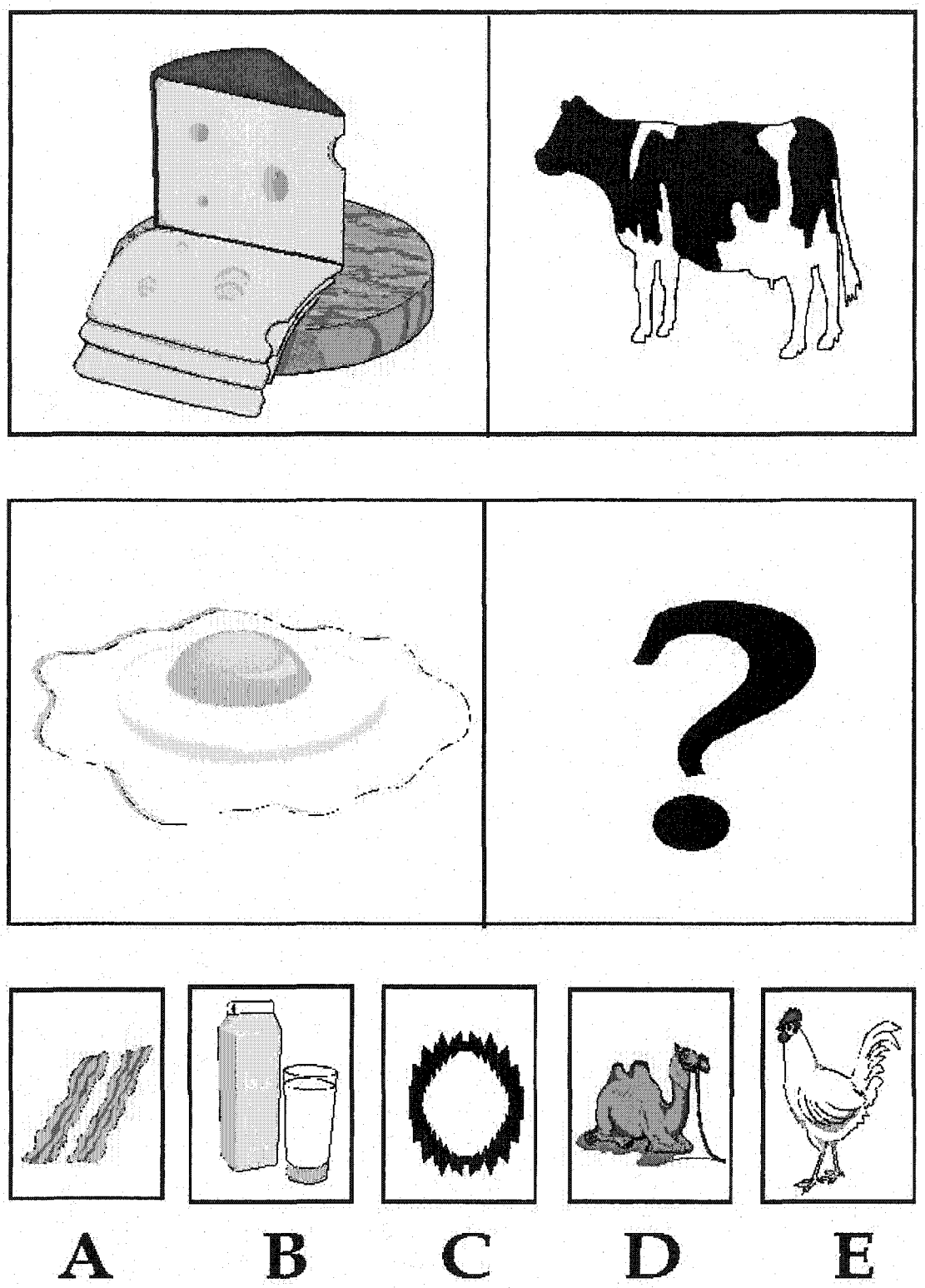
A	B	C	D	E
BABY	TAIL	CHICKEN	PEN	DART

GAT (Version 2) Set A Question 23

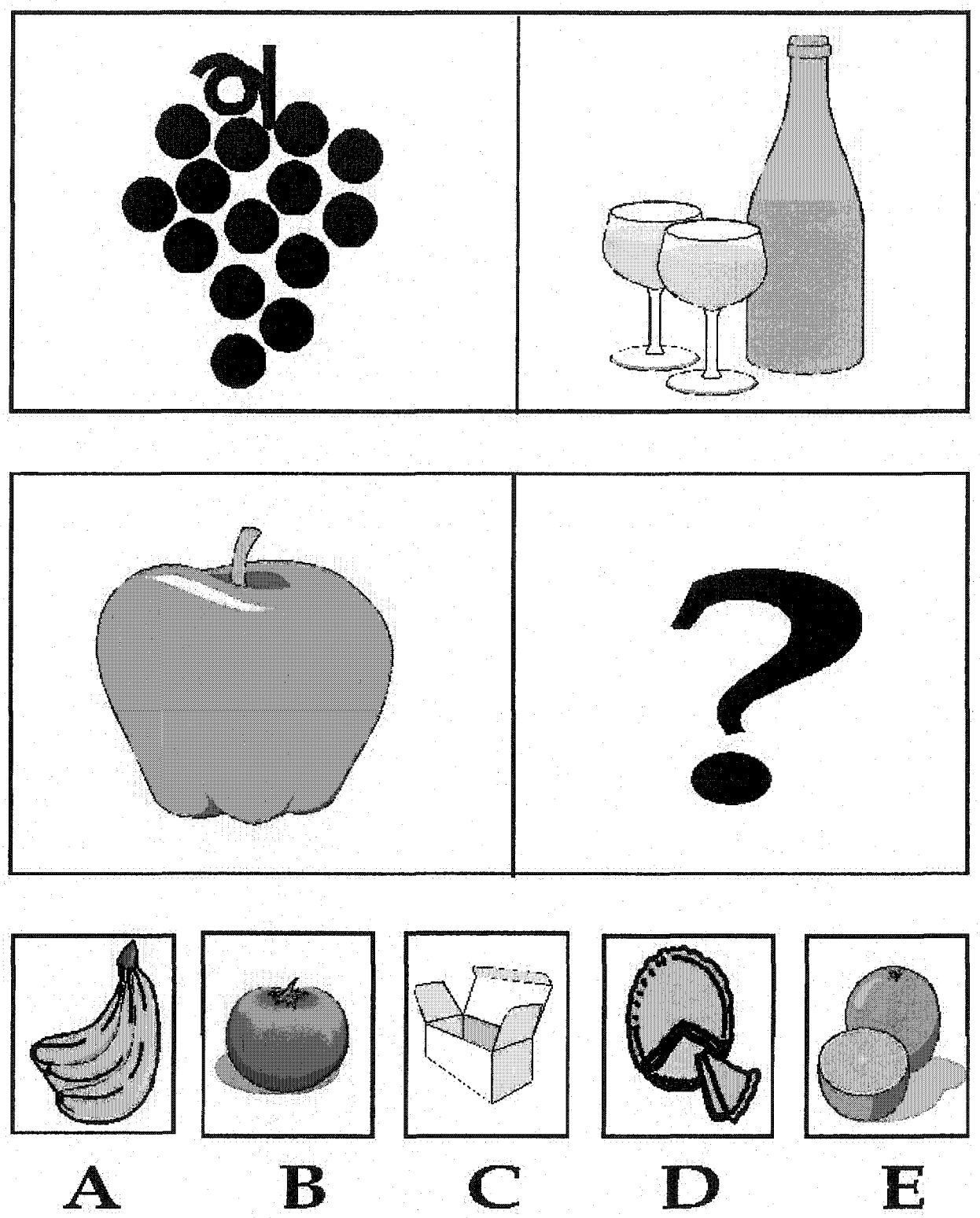


A	B	C	D	E
HATCHET	HANDCUFFS	GLASS	MEAT	FORK

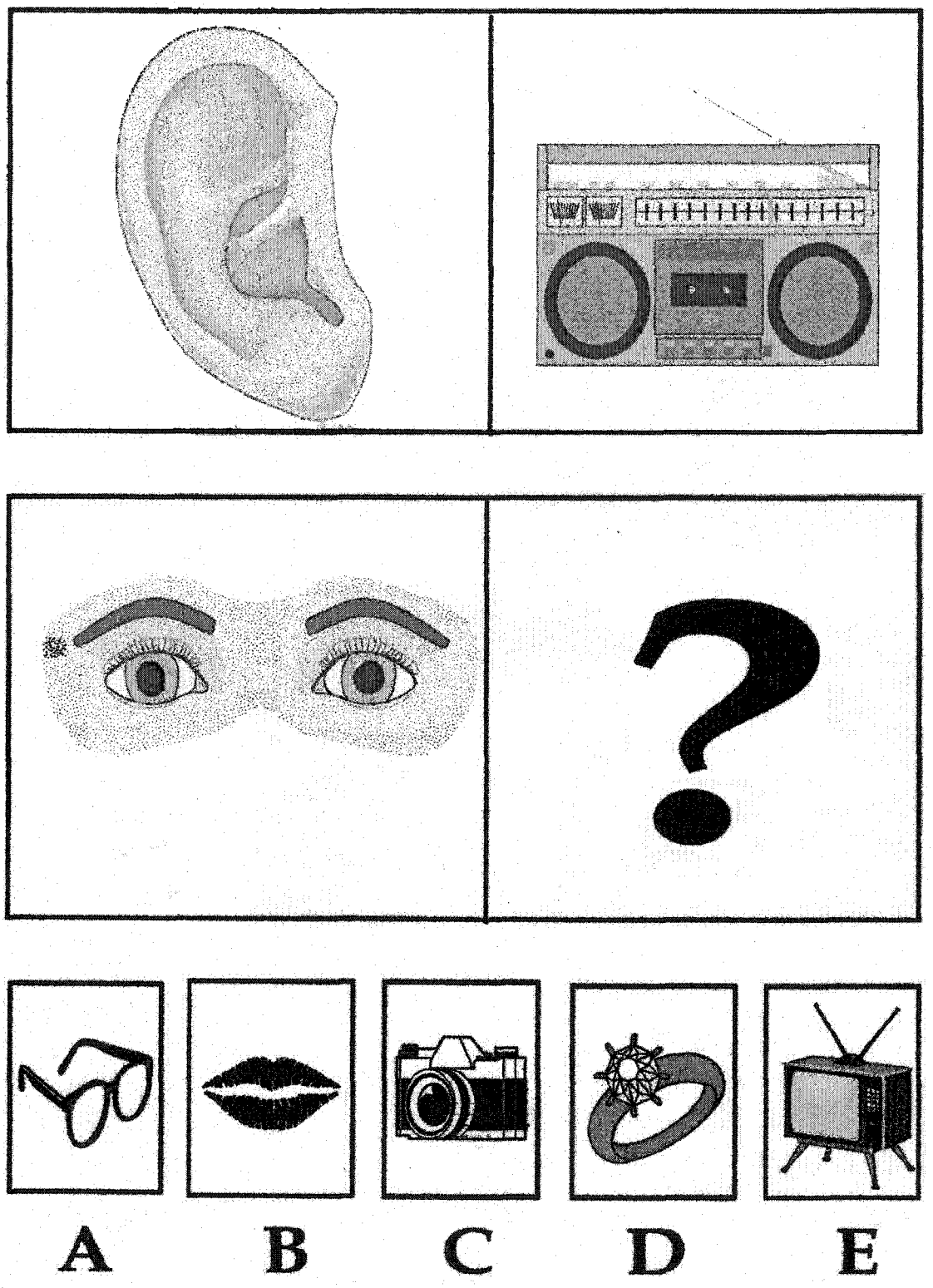
GAT (Version 2) Set A Question 24



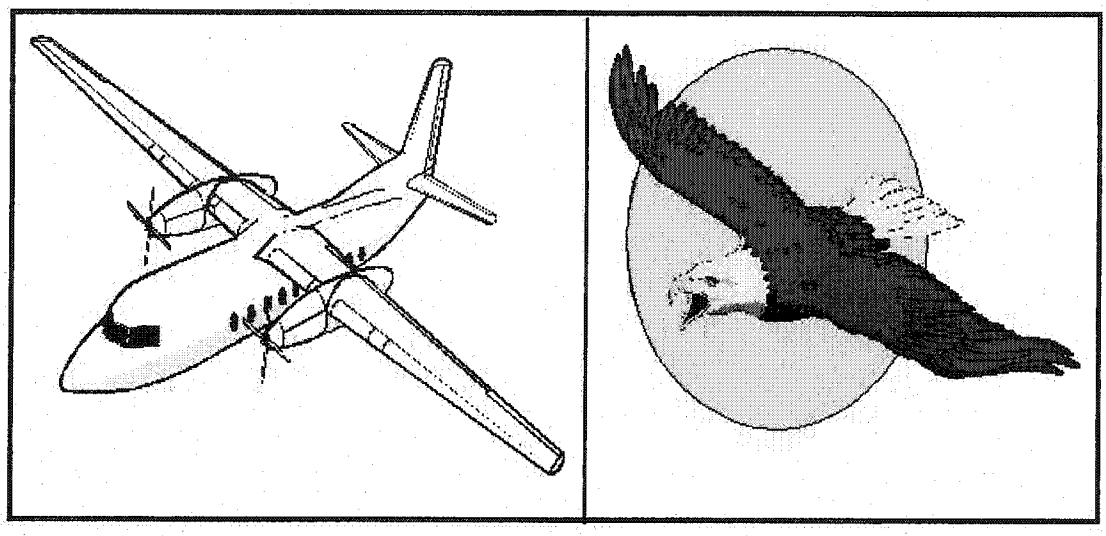
GAT (Version 2) Set B Question 1

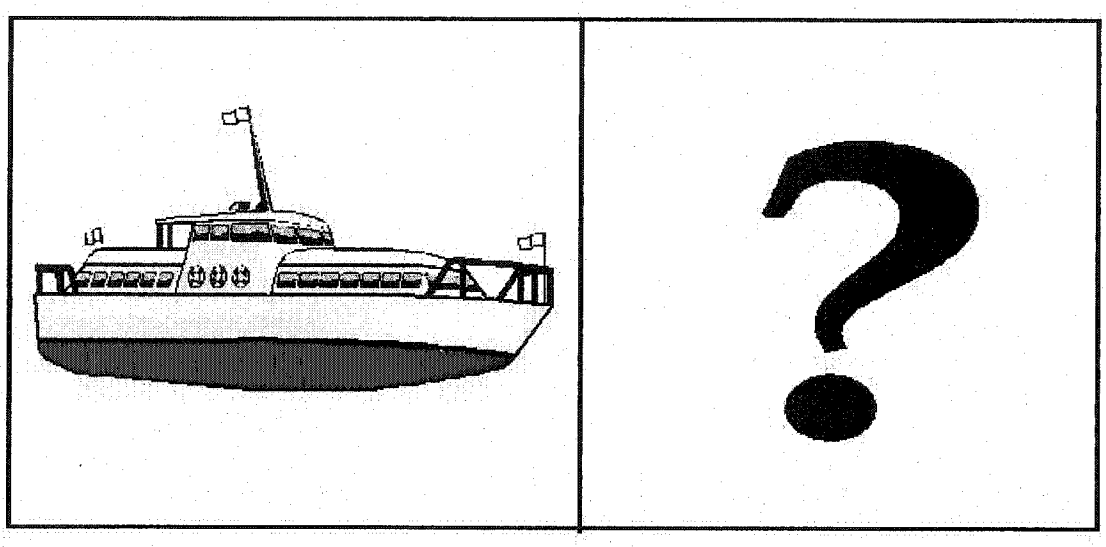


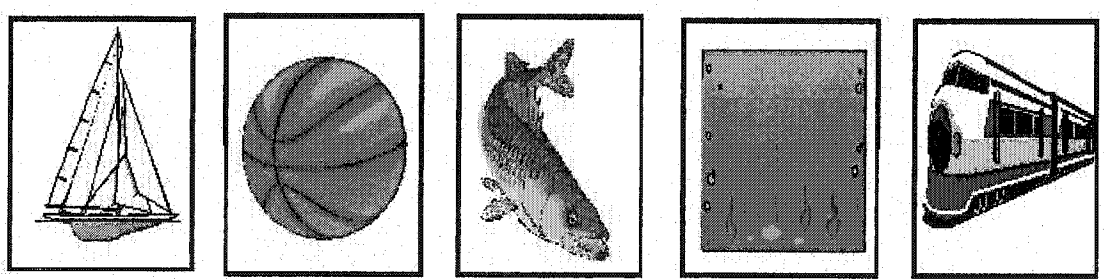
GAT (Version 2) Set B Question 2



GAT (Version 2) Set B Question 3

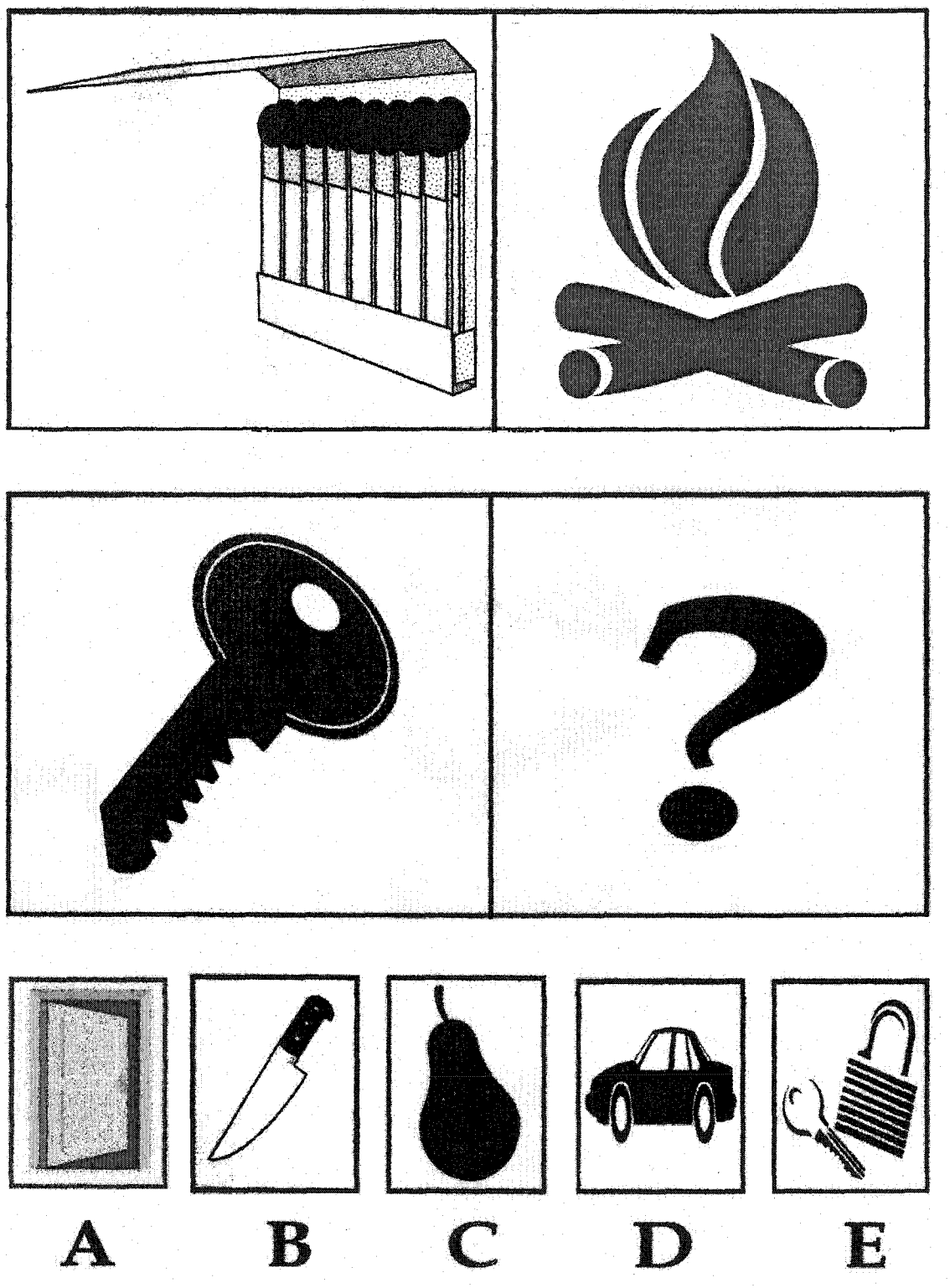




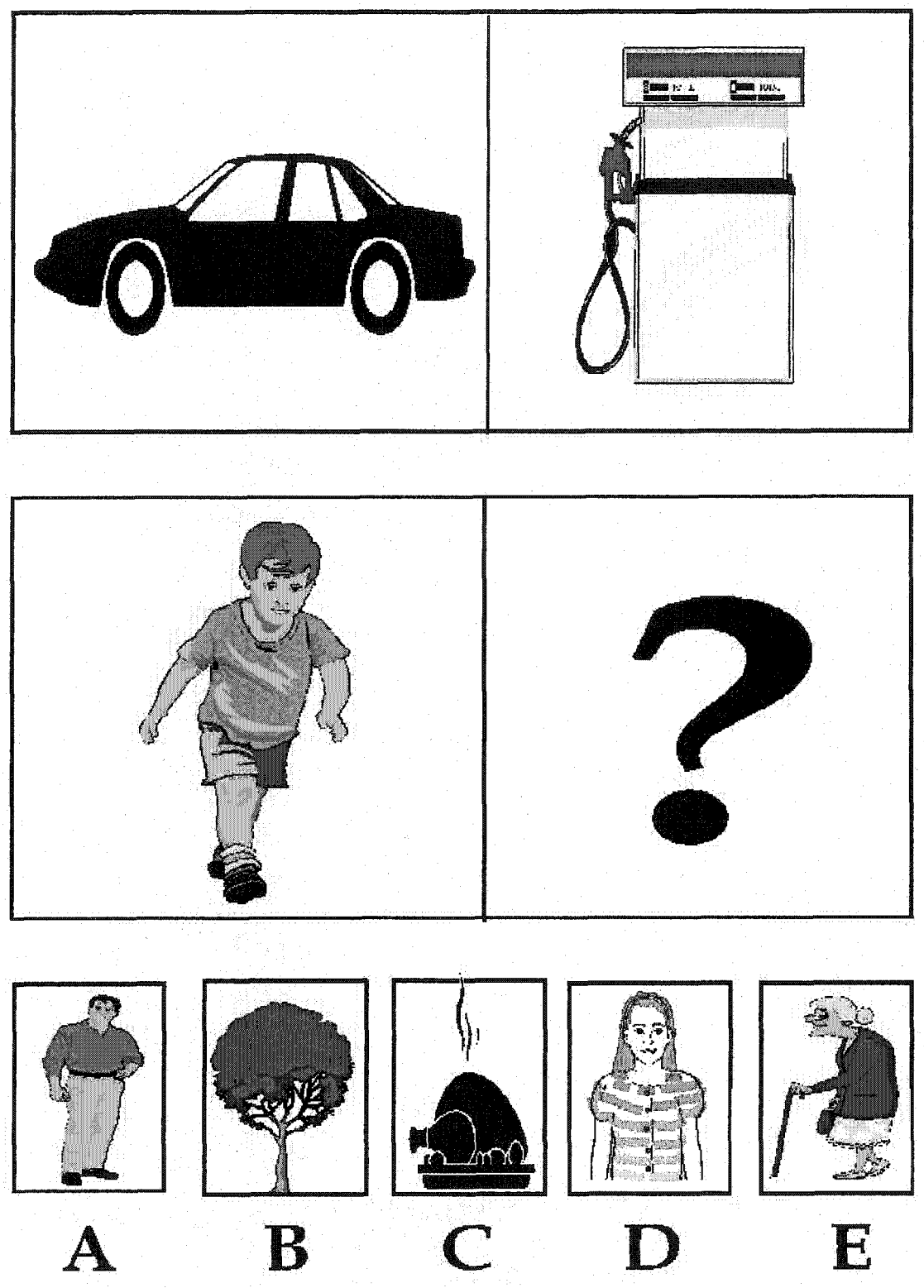


A B C D E

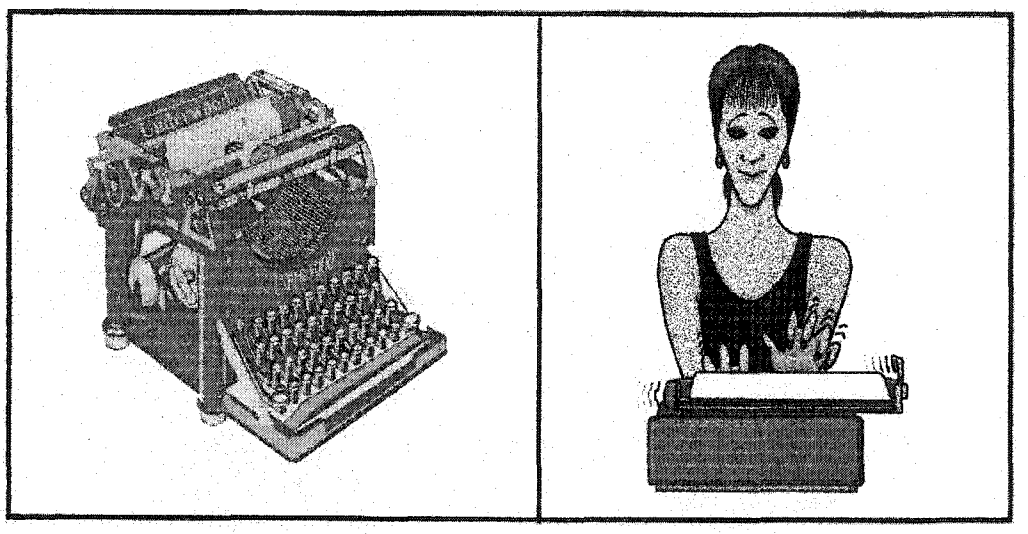
GAT (Version 2) Set B Question 4

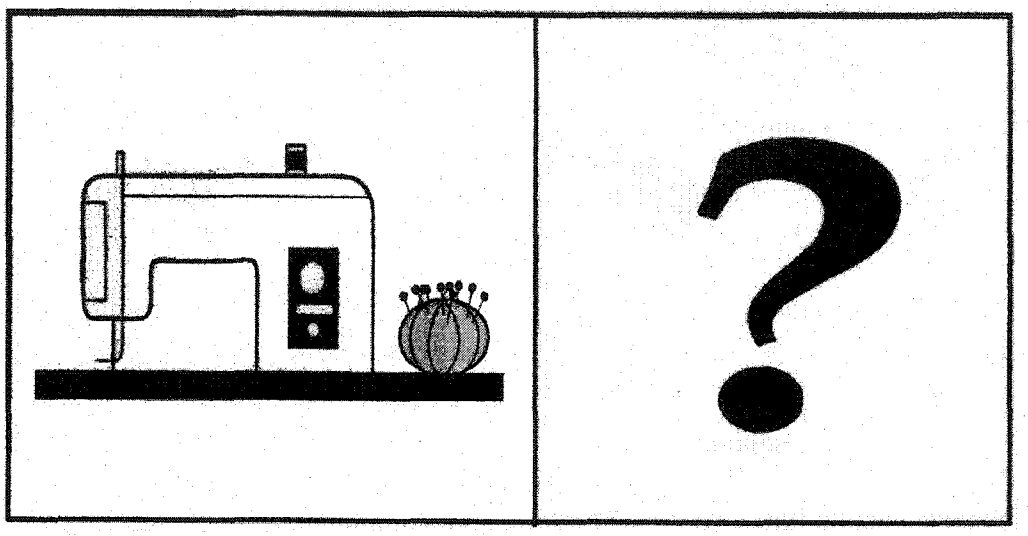


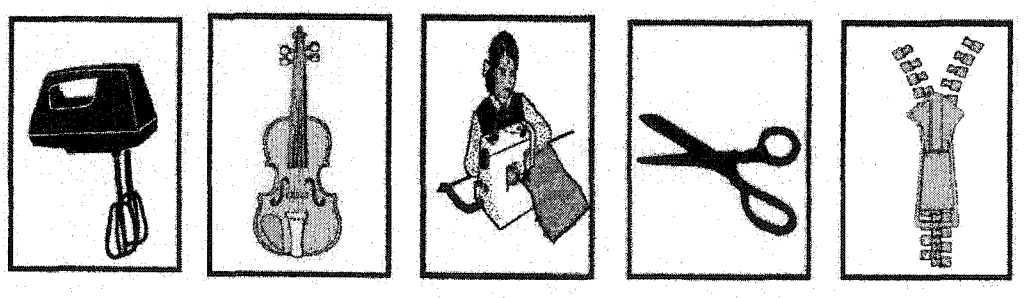
GAT (Version 2) Set B Question 5



GAT (Version 2) Set B Question 6

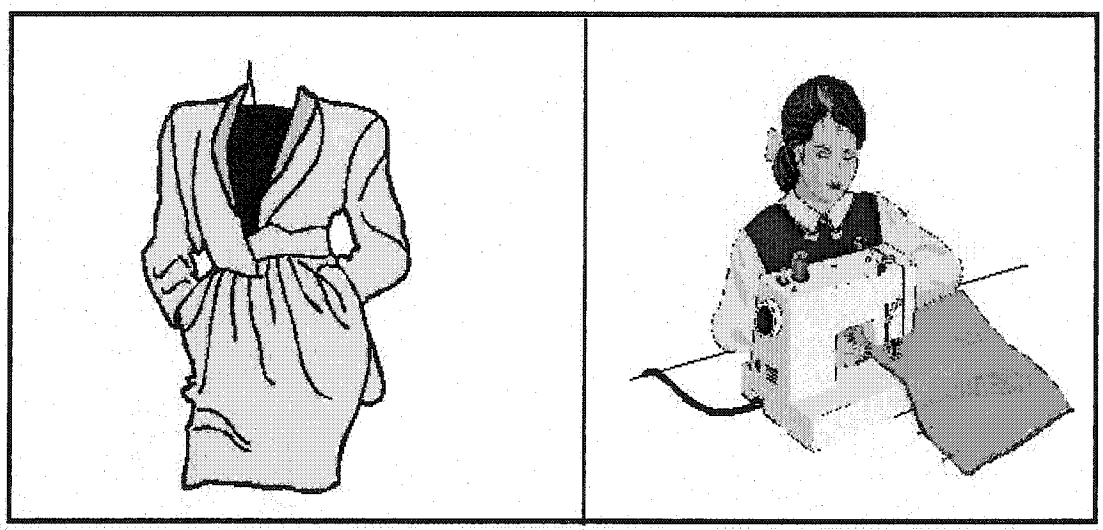


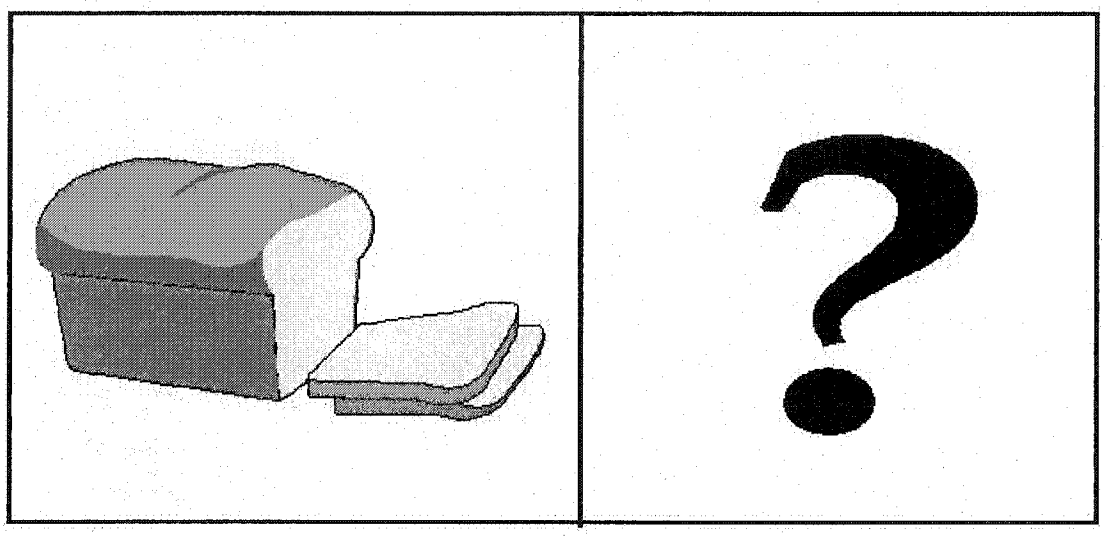


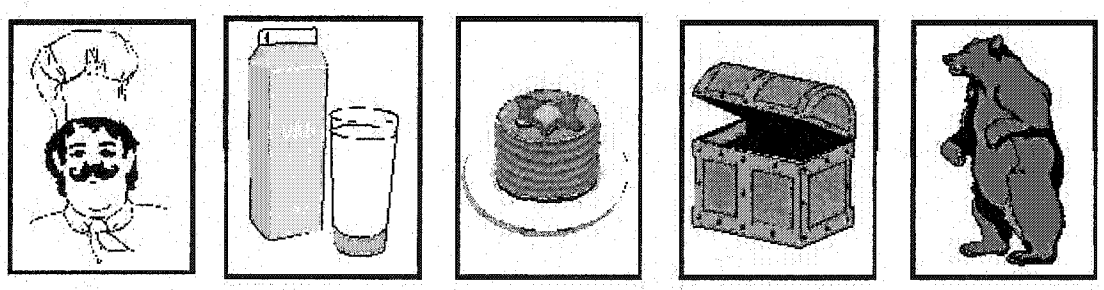


A B C D E

GAT (Version 2) Set B Question 7

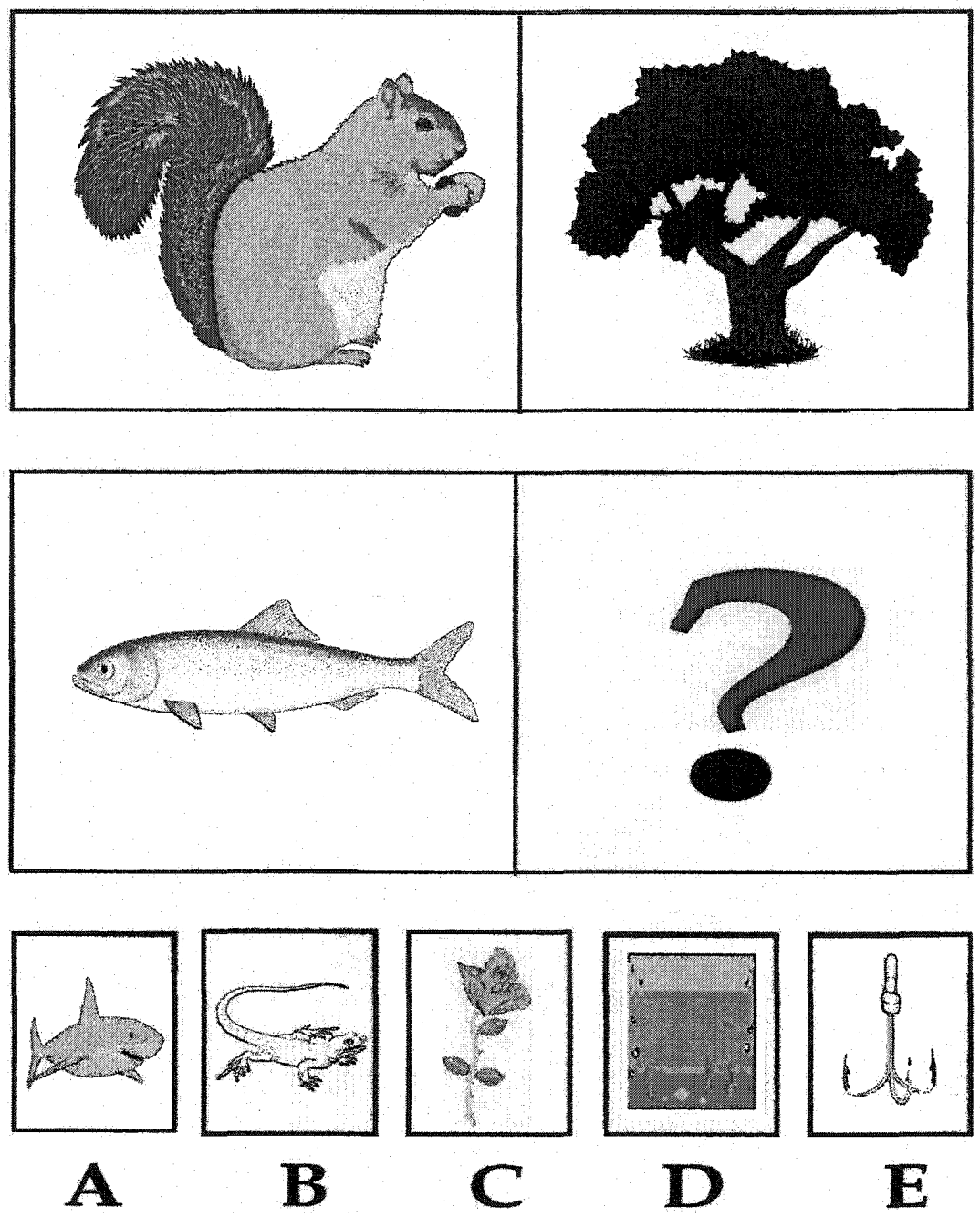




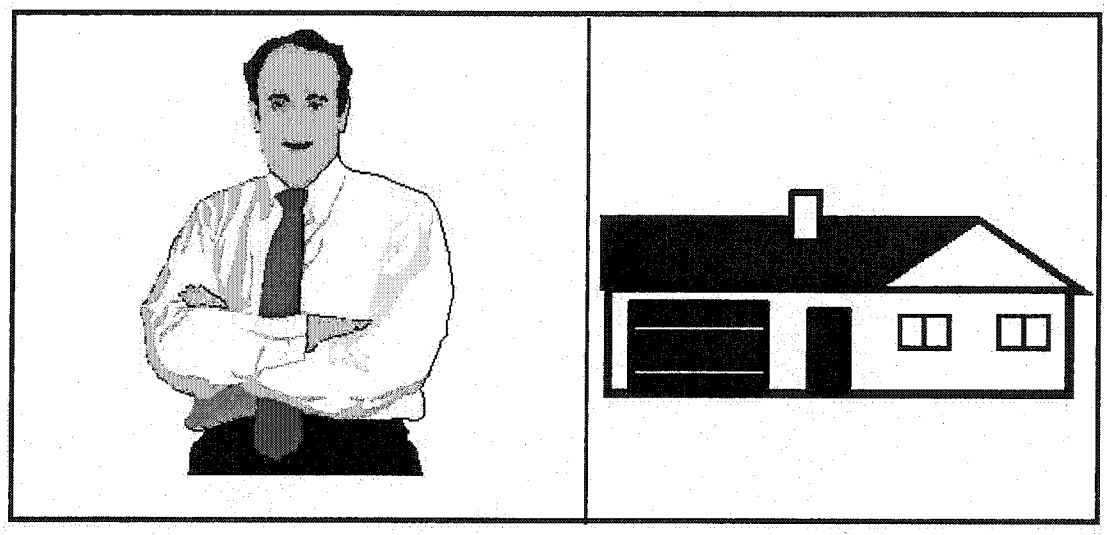


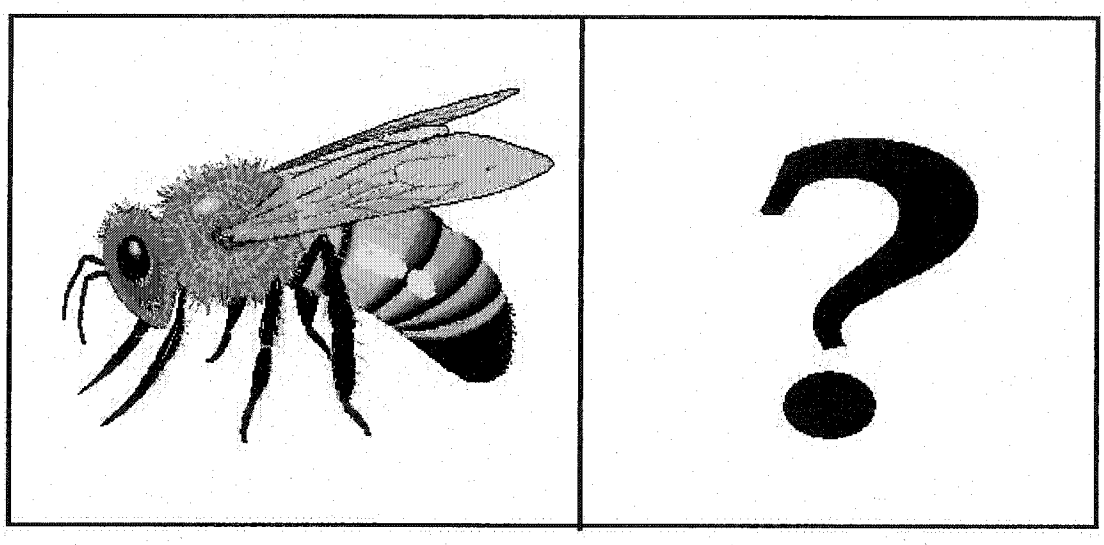
A B C D E

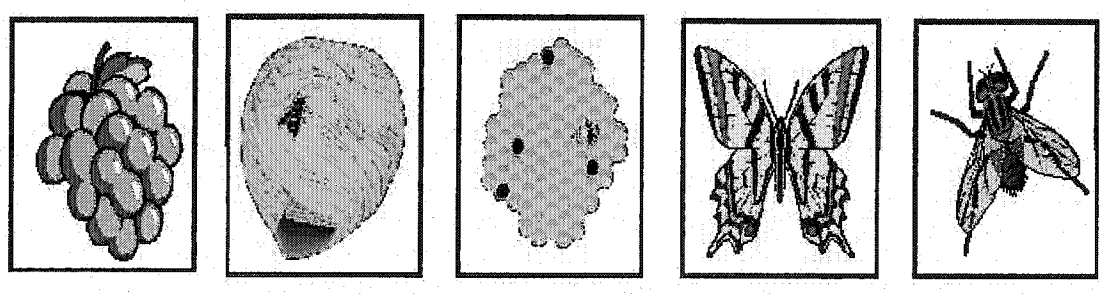
GAT (Version 2) Set B Question 8



GAT (Version 2) Set B Question 9

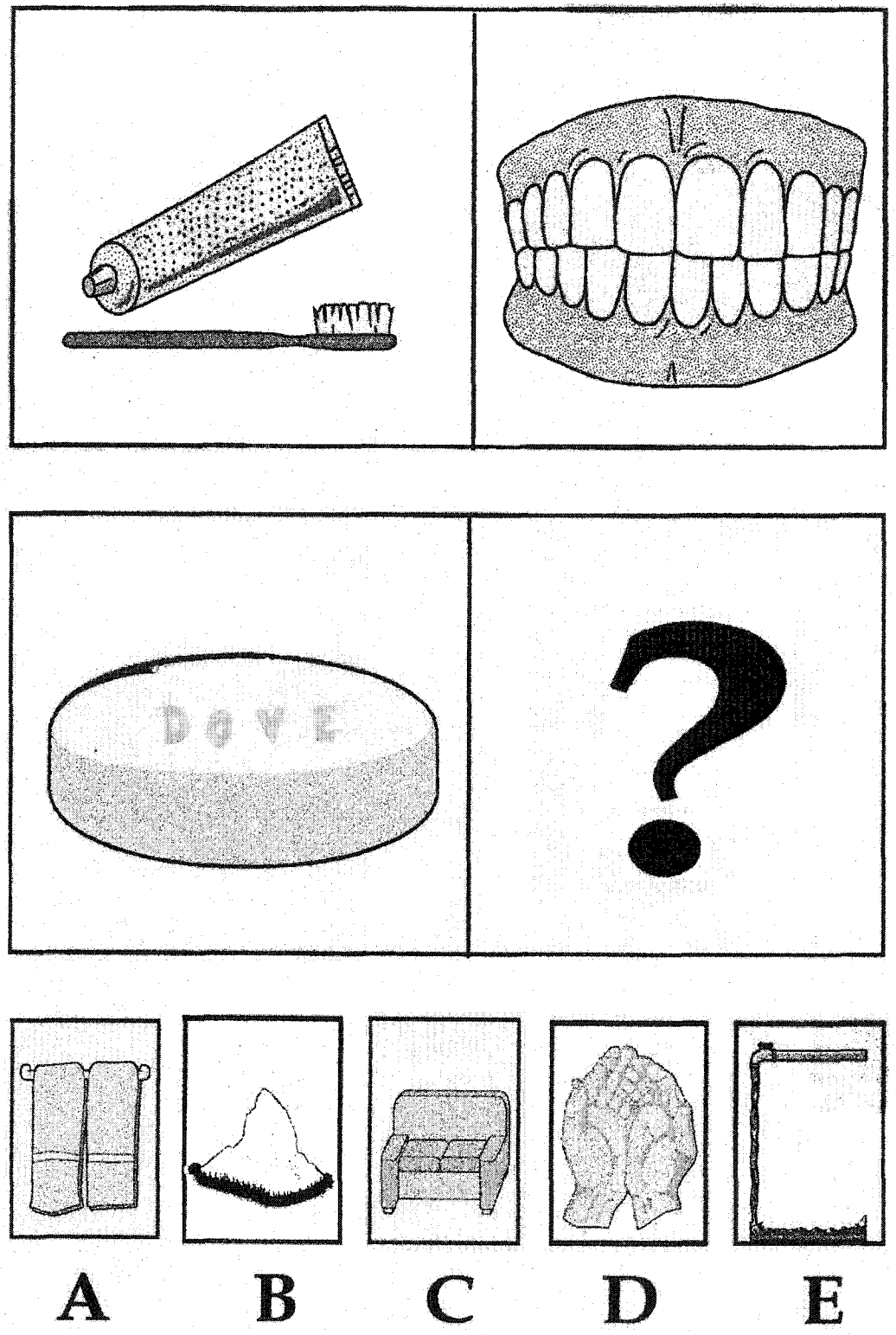




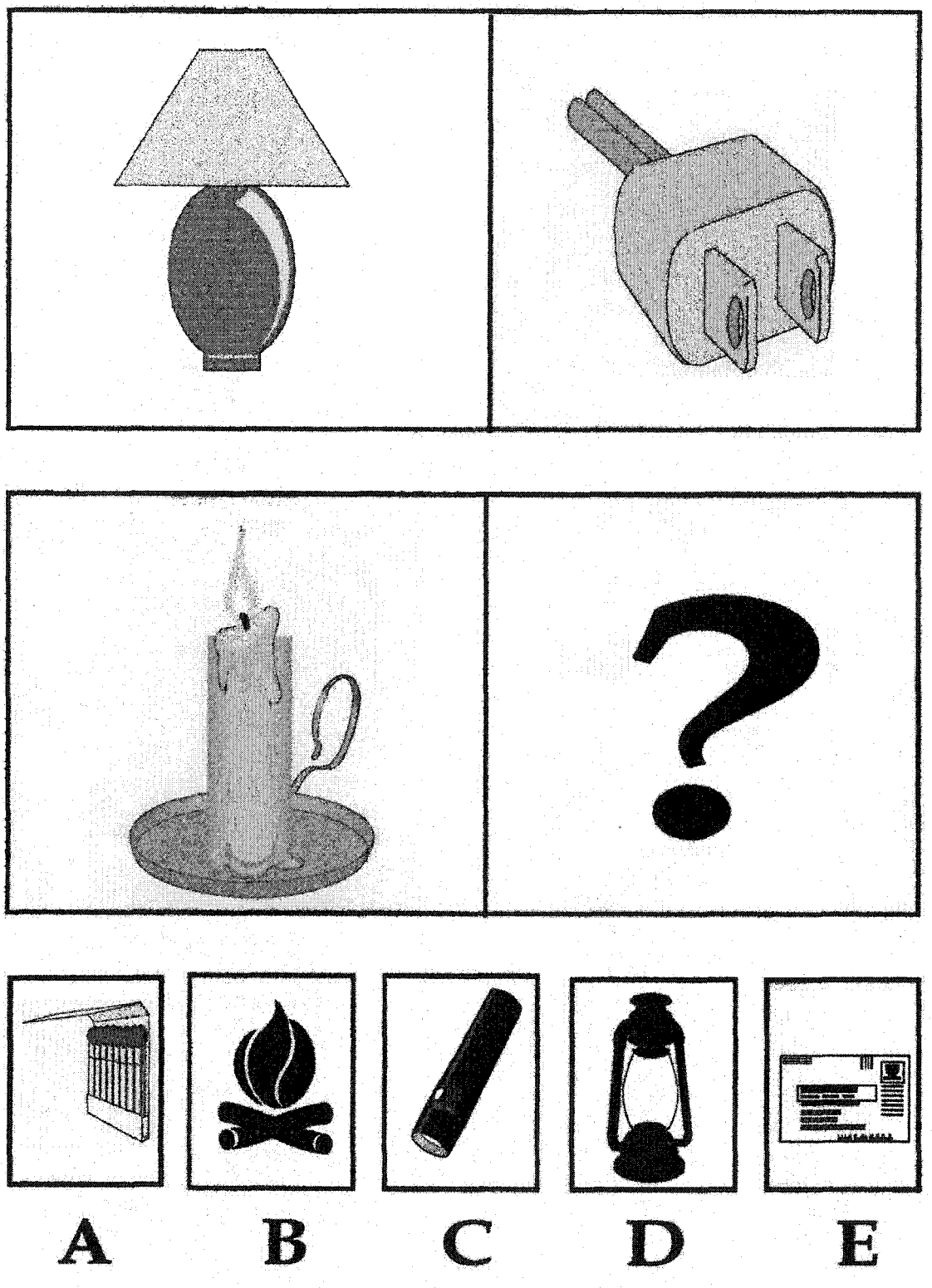


A B C D E

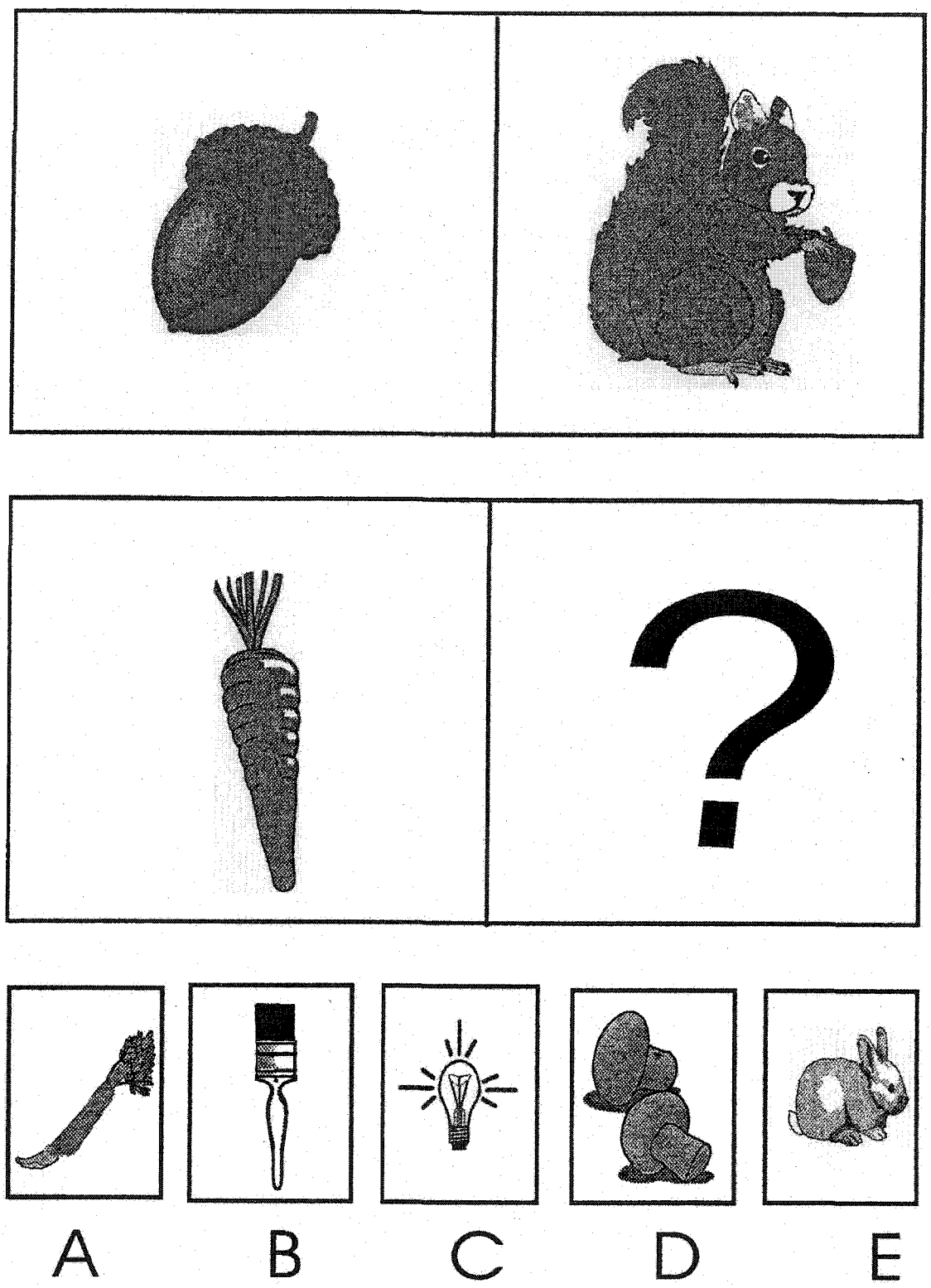
GAT (Version 2) Set B Question 10



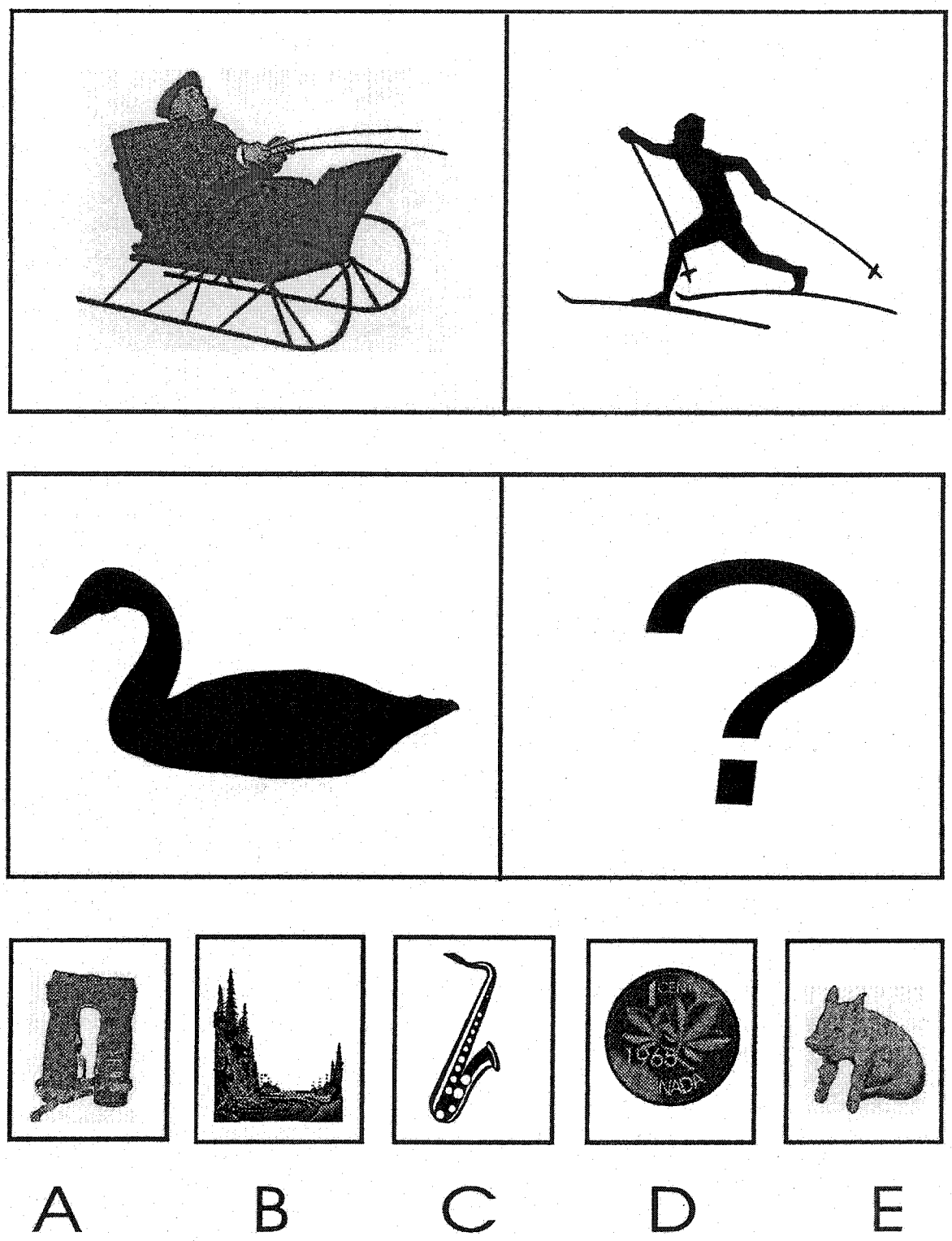
GAT (Version 2) Set B Question 11



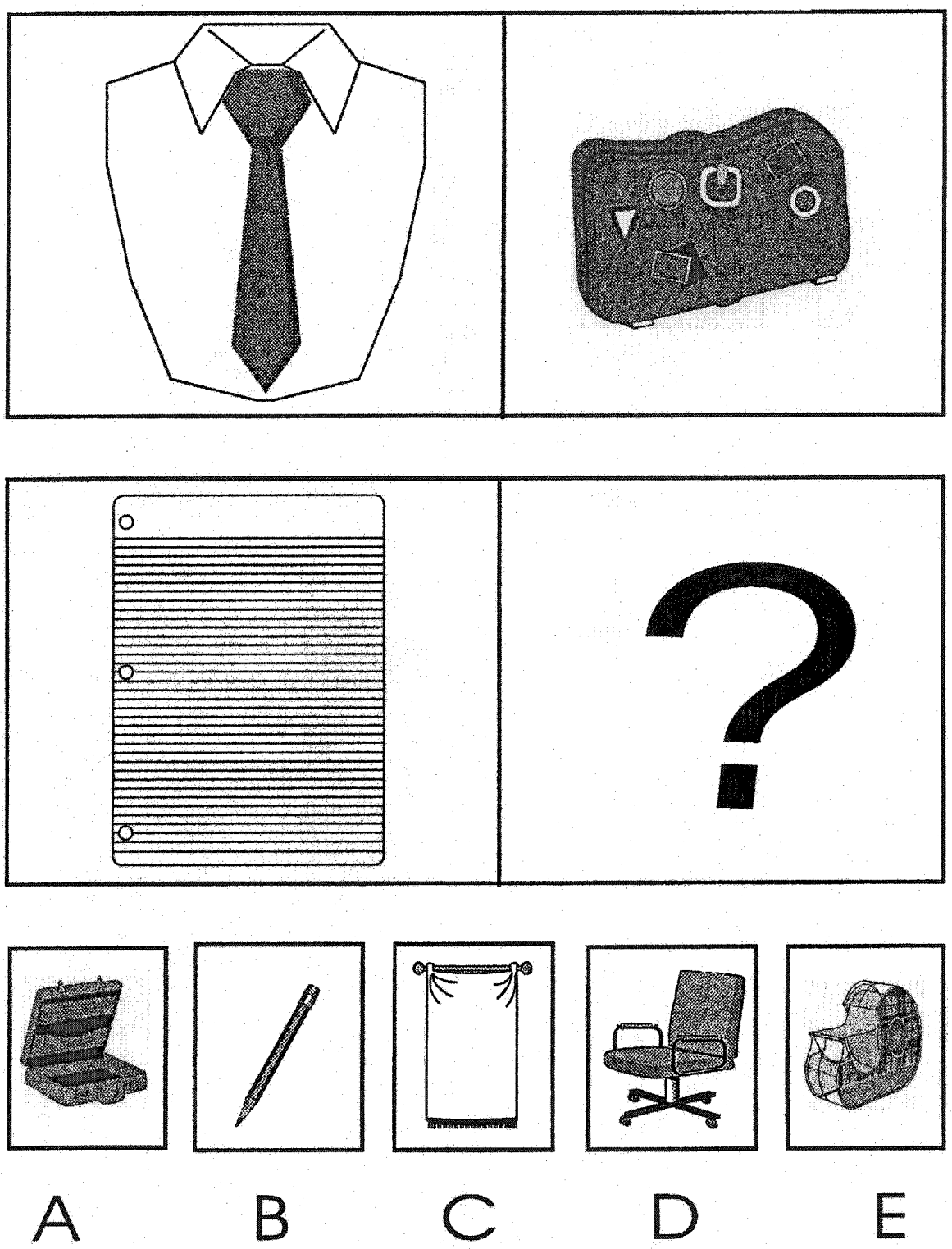
GAT (Version 2) Set B Question 12



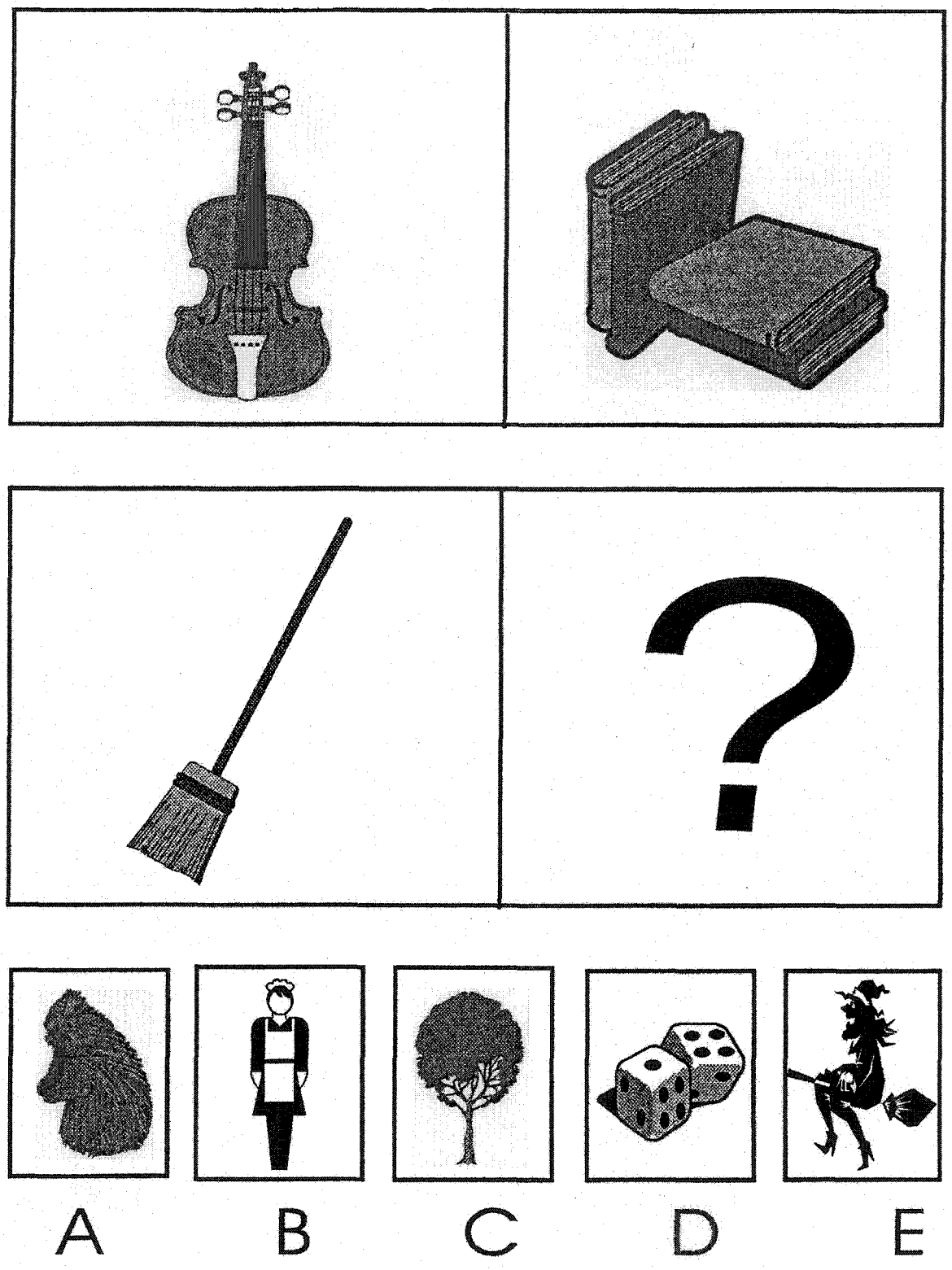
GAT (Version 2) Set B Question 13



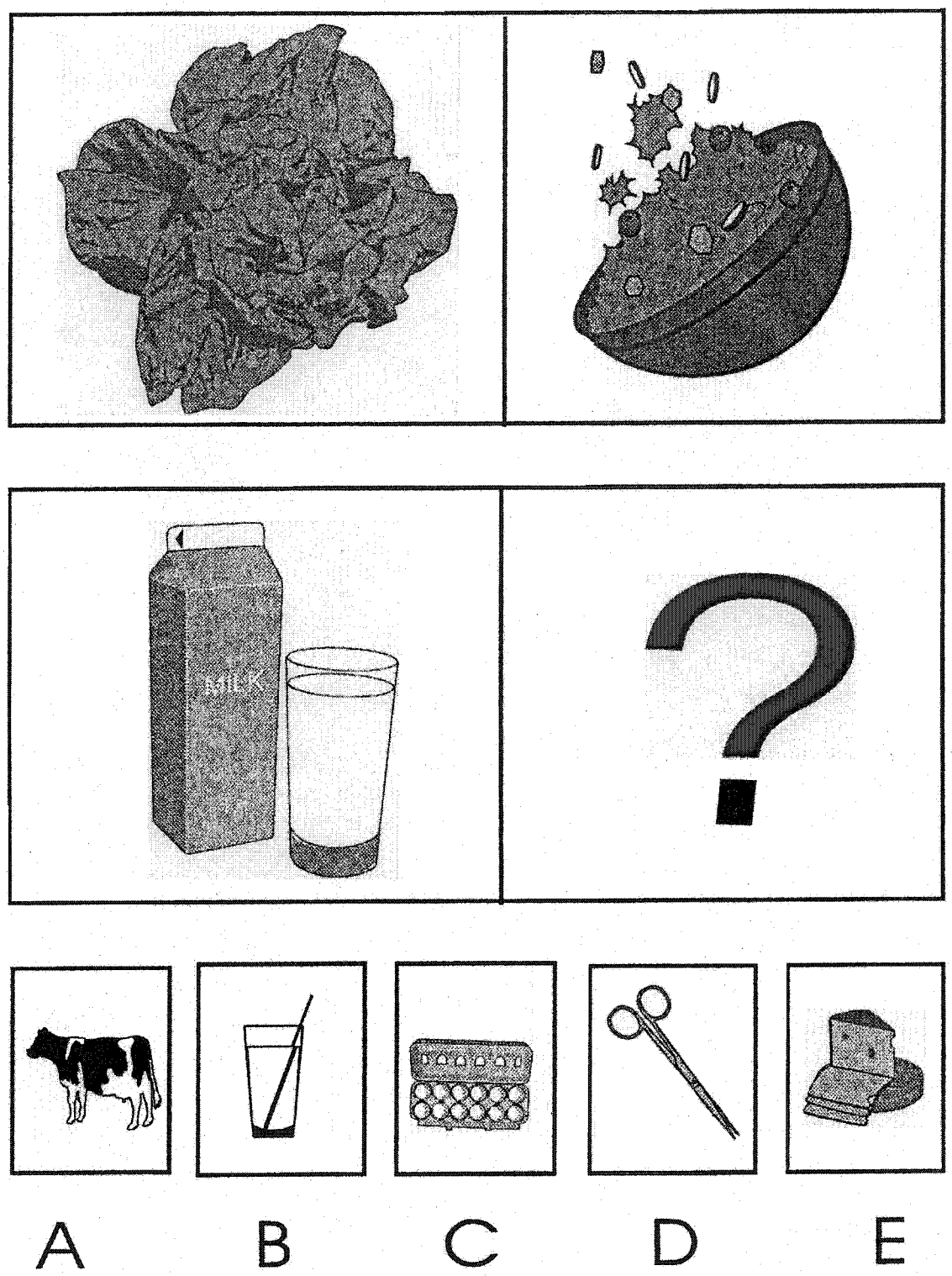
GAT (Version 2) Set B Question 14



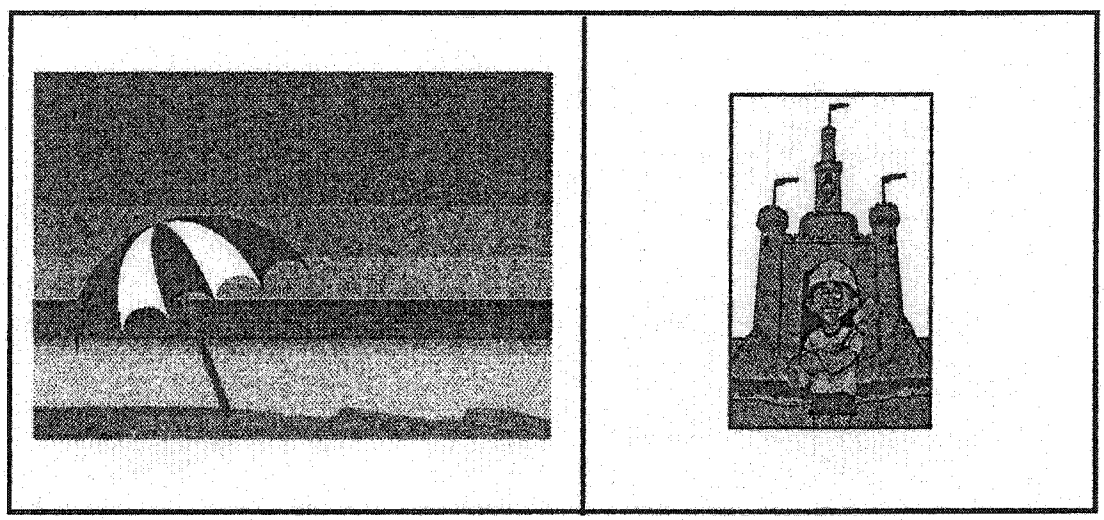
GAT (Version 2) Set B Question 15

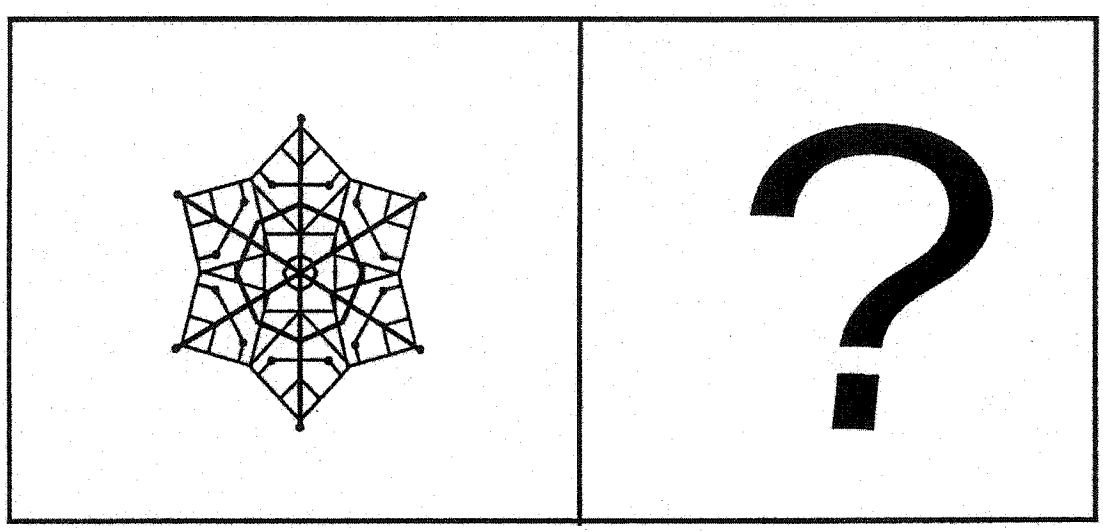


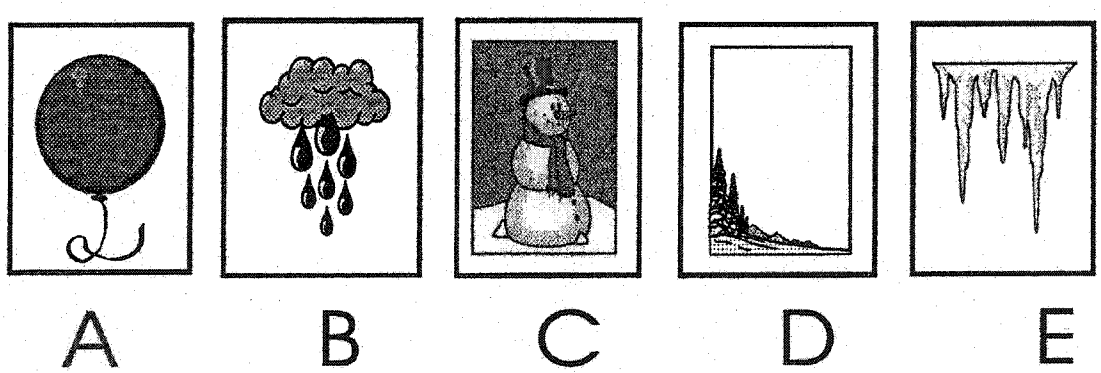
GAT (Version 2) Set B Question 16



GAT (Version 2) Set B Question 17

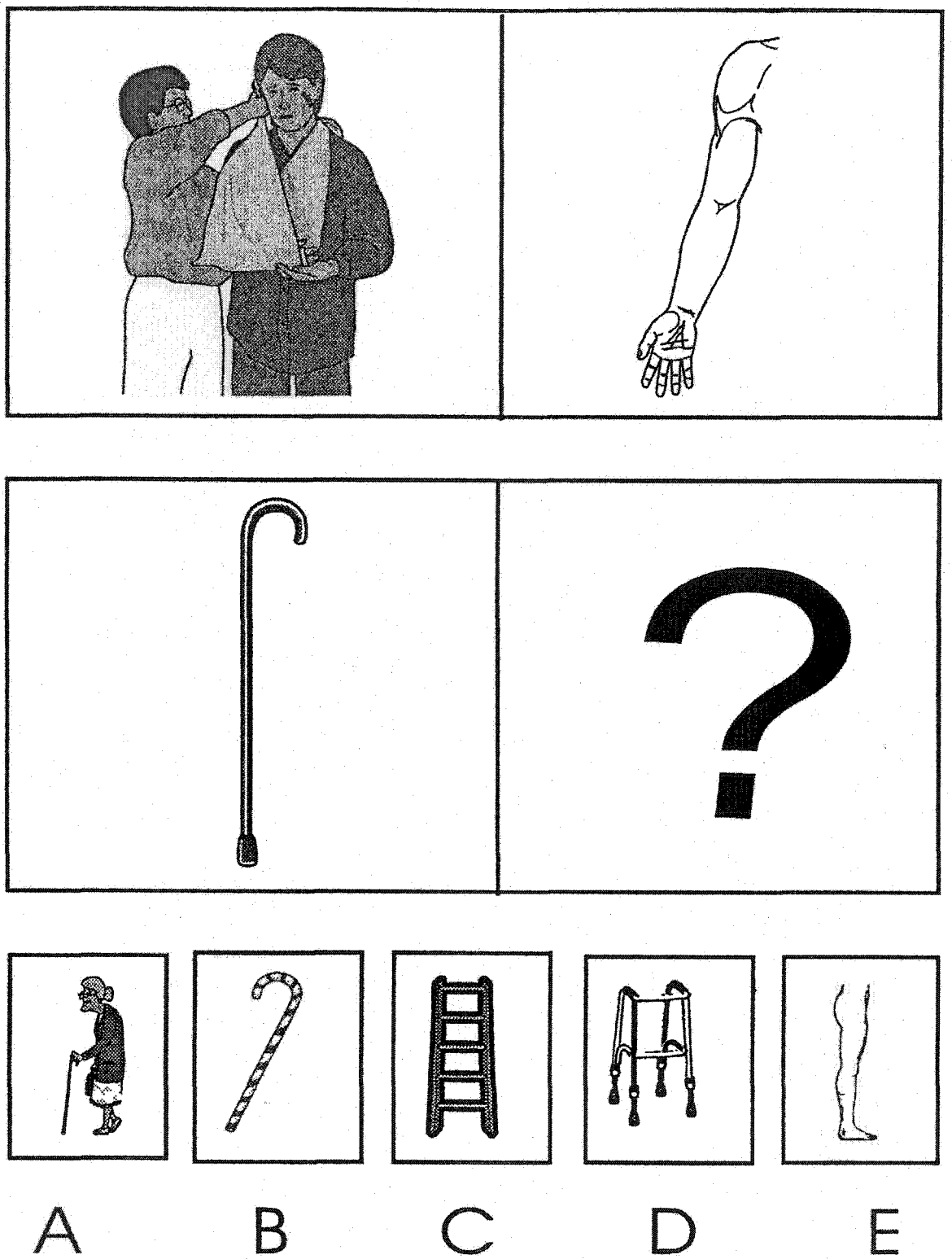




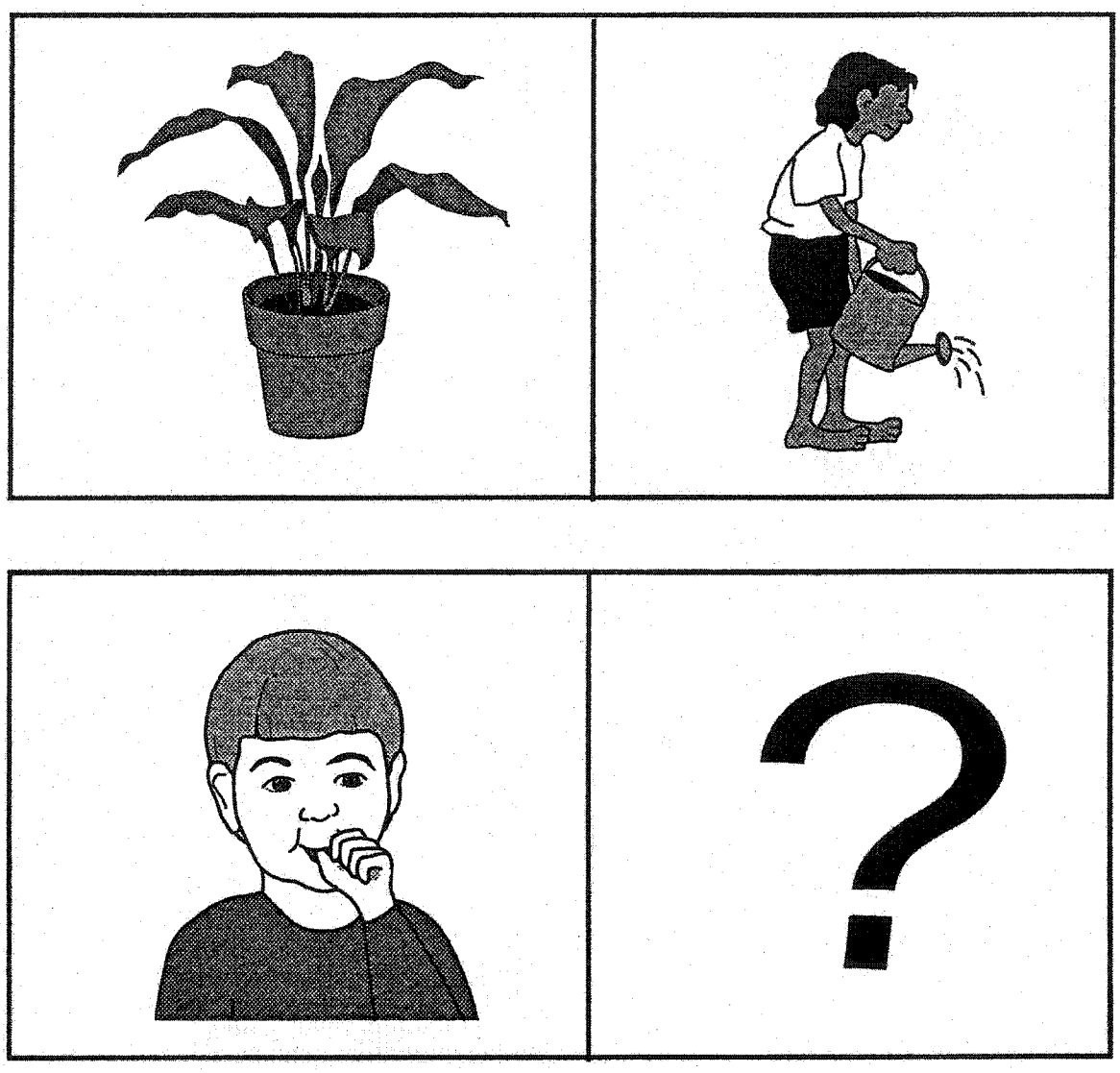


A B C D E

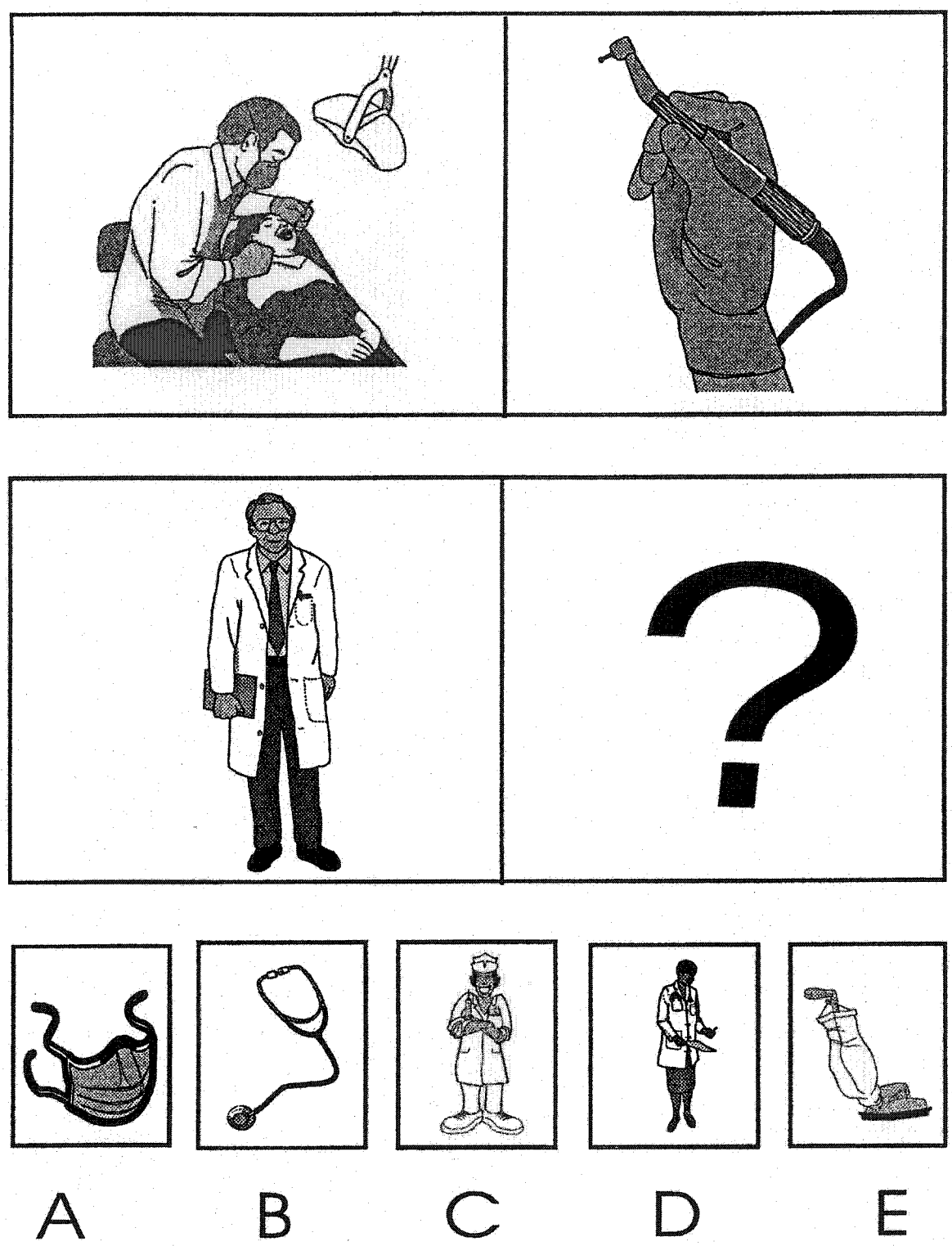
GAT (Version 2) Set B Question 18



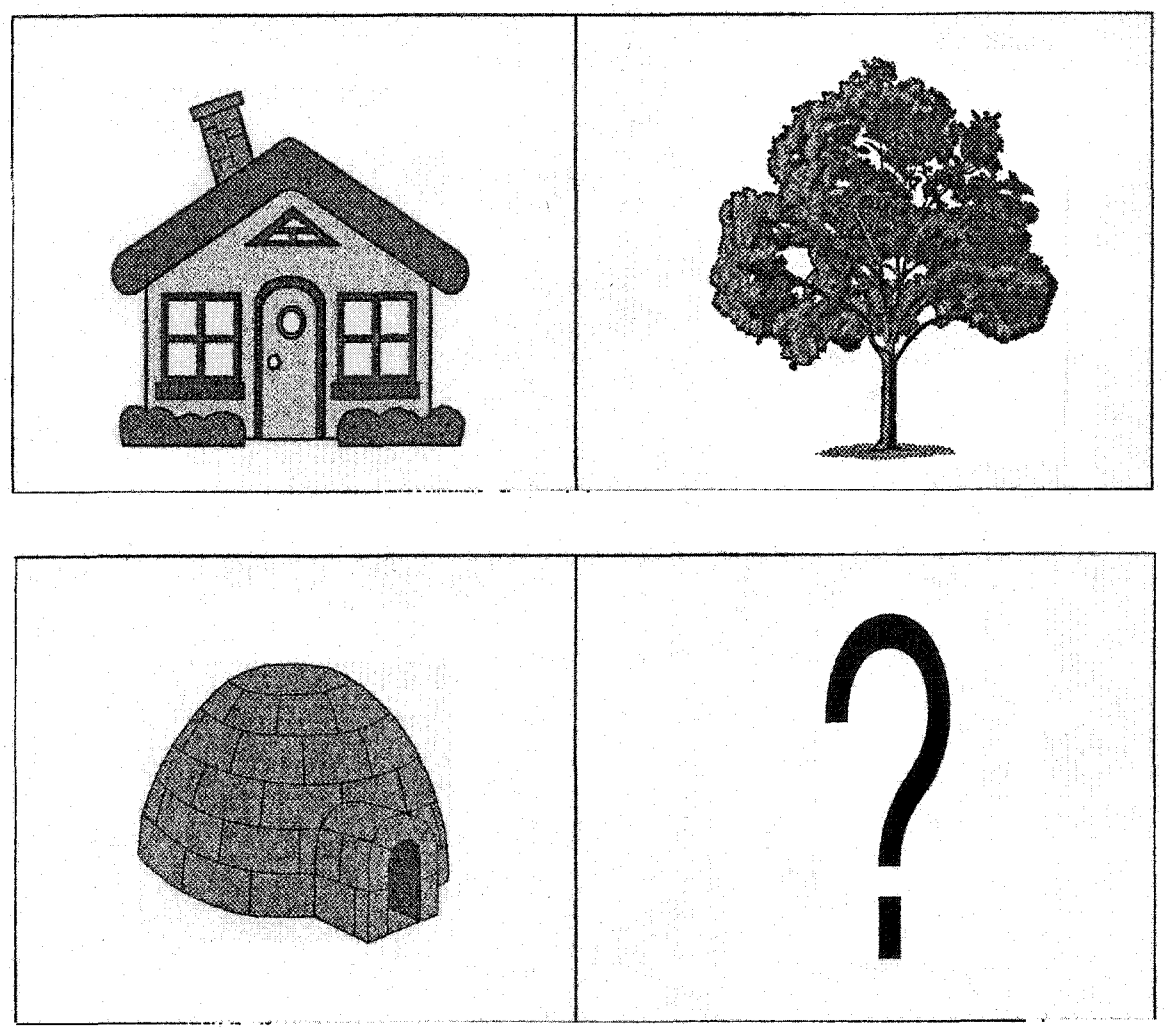
GAT (Version 2) Set B Question 19



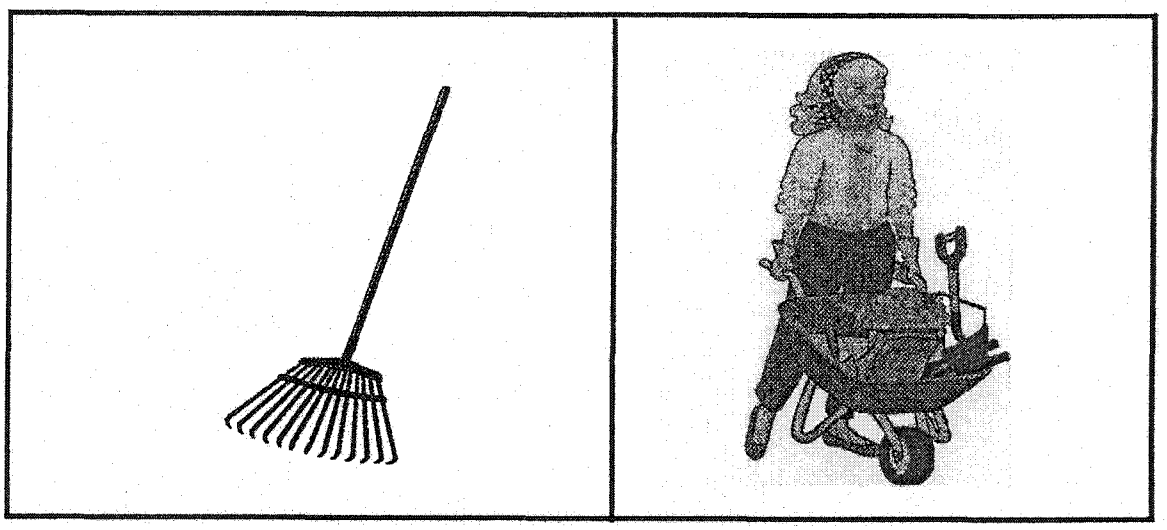
GAT (Version 2) Set B Question 20

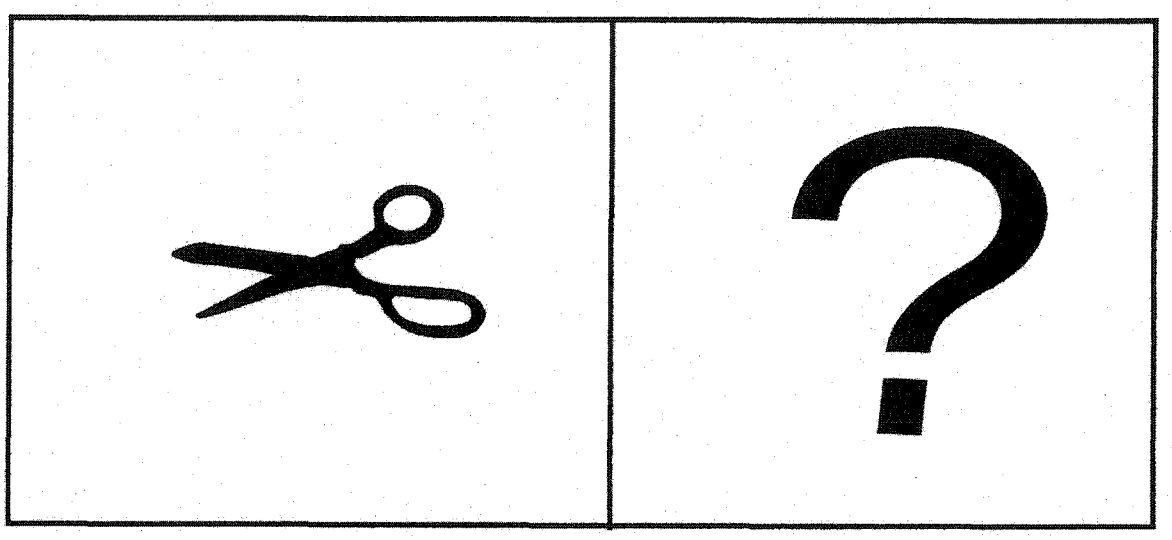


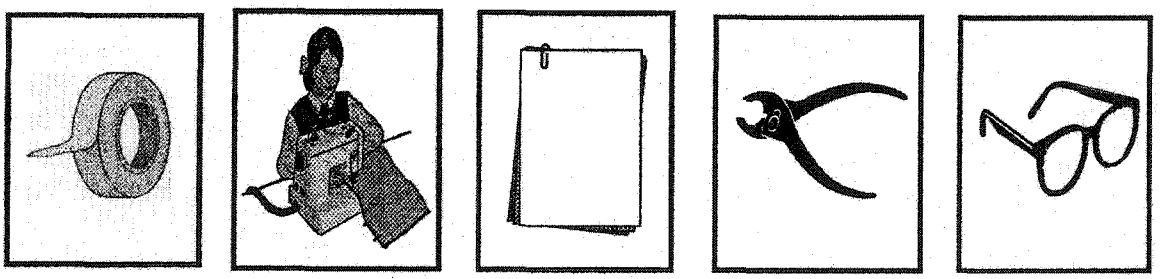
GAT (Version 2) Set B Question 21



GAT (Version 2) Set B Question 22

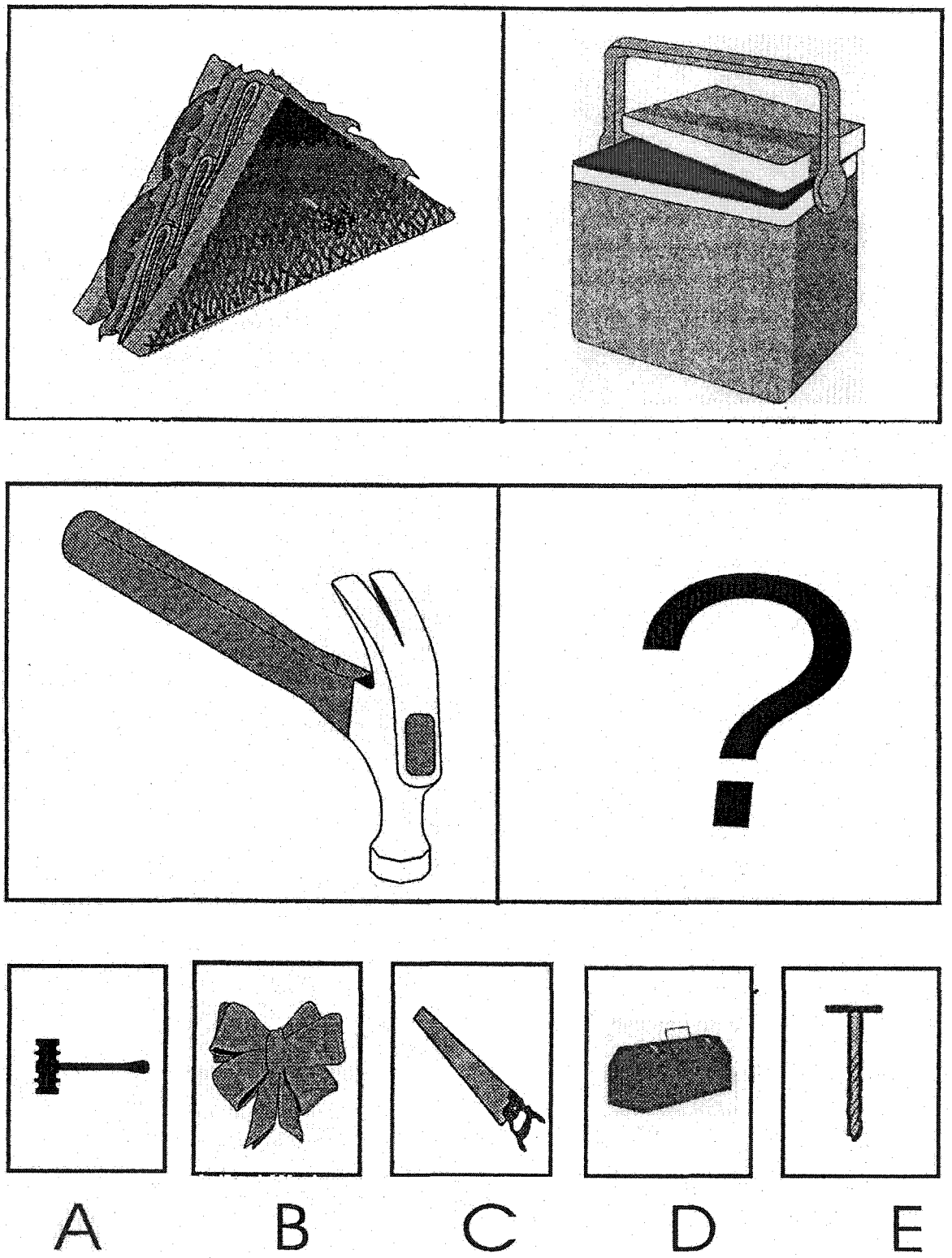






A B C D E

GAT (Version 2) Set B Question 23



GAT (Version 2) Set B Question 24

APPENDIX C

GAT (Versions 1 & 2) Error Classifications

Set A Problems (Words and Pictures):

Question 1: A = associate B = associate C = appearance D = miscellaneous E = correct

Question 2: A = miscellaneous B = correct C = category D = category E = appearance

Question 3: A = appearance B = miscellaneous C = correct D = associate E = category

Question 4: A = miscellaneous B = correct C = category D = associate E = appearance

Question 5: A = category B = associate C = miscellaneous D = correct E = category

Question 6: A = associate B = correct (words)/appearance (pictures) C = correct
D = miscellaneous E = correct

Question 7: A = miscellaneous B = correct C = associate D = category E = appearance

Question 8: A =category B = appearance C = miscellaneous D = correct E = associate

Question 9: A = miscellaneous B = correct C = associate D = category E = appearance

Question 10: A = correct B = associate C = correct D = appearance E = miscellaneous

Question 11: A = associate B = category C = appearance D = miscellaneous E = correct

Question 12: A = appearance B = miscellaneous C = correct D = associate E = category

Question 13: A = category B = correct C = associate D = appearance E = miscellaneous

Question 14: A = correct B = correct C = appearance D = miscellaneous E = category

Question 15: A = miscellaneous B = category C = correct D = associate E = appearance

Question 16: A = miscellaneous B = category C = correct D = associate E = appearance

Question 17: A = appearance B = miscellaneous C = category D = correct E = associate

Question 18: A =correct B = associate C = category D = miscellaneous E = appearance

Question 19: A= associate B = correct C = associate D = appearance E = miscellaneous

Question 20: A= category B = appearance C = miscellaneous D = associate E = correct

Question 21: A= miscellaneous B= category C= correct D= associate E= appearance

Question 22: A = category B = associate C = correct D = appearance E = miscellaneous

Question 23: A = miscellaneous B = category C = correct D = associate E = appearance

Question 24: A = appearance B = miscellaneous C = category D = correct E = associate

---

Set B Problems (Words and Pictures):

Question 1: A = associate B = category C = appearance D = miscellaneous E = correct

Question 2: A = category B = appearance C = miscellaneous D = correct E = associate

Question 3: A = associate B = category C = appearance D = miscellaneous E = correct

Question 4: A = appearance B = miscellaneous C = correct D = associate E = category

Question 5: A = category B = appearance C = miscellaneous D = correct E = associate

Question 6: A = category B = miscellaneous C = correct D = associate E = appearance

Question 7: A = appearance B = miscellaneous C = correct D = associate E = category

Question 8: A = correct B = associate C = category D = appearance E = miscellaneous

Question 9: A = category B = appearance C = miscellaneous D = correct E = associate

Question 10: A = miscellaneous B = correct C = associate D = category E = appearance

Question 11: A = category B = appearance C = miscellaneous D = correct E = associate

Question 12: A = correct B = associate C = category D = appearance E = miscellaneous

Question 13: A = category B = appearance C = miscellaneous D = category E = correct

Question 14: A = correct B = associate C = appearance D = miscellaneous E = category

Question 15: A = correct B = associate C = appearance D = miscellaneous E = category

Question 16: A = appearance B = associate C = miscellaneous D = correct E = associate

Question 17: A = associate B = appearance C = miscellaneous D = category E = correct

Question 18: A = miscellaneous B = category C = correct D = associate E = appearance

Question 19: A = associate B = appearance C = miscellaneous D = category E = correct

Question 20: A = correct B = associate C = appearance D = miscellaneous E = category

Question 21: A = category B = correct C = associate D = appearance E = miscellaneous

Question 22: A = appearance B = miscellaneous C = category D = correct E = associate

Question 23: A = category B = correct C = associate D = appearance E = miscellaneous

Question 24: A = appearance B = miscellaneous C = category D = correct E = associate

APPENDIX D

GAT Administration Instructions

Testing Conditions

The GAT, like any other psychological test, should be administered in a quiet room free from distracting interruptions. The test was designed to be administered to individual examinees rather than in a group setting. The GAT can be administered in less than 20 minutes to *cognitively intact* older adults. Dementia patients require longer (range: 25 - 40 minutes). There are no discontinuation criteria. Examinees should be encouraged to guess even if unsure of the answer. If an examinee is aware that he or she is doing poorly on the test, the examiner is encouraged to offer statements of reassurance such as, "This is a hard task for many people. Just keep trying your best."

Administration

Present the first practice item in picture format. Begin by saying: **This is a test of your ability to think about the relationships between pictures and words. Look at this problem. This picture goes with this picture (examiner points to A and B terms of the analogy). Can you think of one reason why these two pictures go together? (Examiner provides correct answer if examinee is wrong and spends some time coaching/teaching the rationale). That's right. Just like the bird lives in the nest, the dog (examiner points to the C term) lives where? Which of these answer choice options down here provides the best answer? Why? (examiner coaches if not correct and provides the correct answer).**

Good. Now look at this word problem. This word goes with this word (examiner points to A and B terms of the analogy). Why? (Examiner provides correct answer if examinee is wrong and teaches the rationale). So. Just like sheep's wool is used to make woolen sweater, a cow's hide is used to make leather boots.

The examiner then turns back to the first practice problem in picture format and says,

"I am going to be showing you some more problems just like the two we just tried in either picture or word format. The instructions will be as follows: 'Just like this (examiner points to bird) goes with this (examiner points to nest), this (examiner points to dog) goes with which of these down here?' (examiner runs finger across answer choice options)

Just like we practiced, once you have figured out the relationship between the first two items, this relationship will help you to determine which answer choice down here is best. Make sure to look at all of the answer choice options because some of the answer choices may be related to this picture here (examiner points to the C term) but they are not correct.

For example, look down her. Notice that this picture of the dog (examiner points to the C term) is related to this picture of the cat (examiner points to relevant answer choice option). Dog and cat are related to each other but "cat" is not the correct answer because... just like a bird lives in a nest, a dog does not "live in" a cat.

Present the first test item.

Now let's try some more. Just like this picture (word) goes with this picture (word), this picture (word) goes with which of these pictures (words) down here?

Do not offer any further coaching. If the examinee is having difficulty selecting an answer choice, encourage him/her to guess. There are no time limits so the examinee may take as long as needed to respond. The examiner may wish to periodically ask why the examinee made the choice he/she did, making sure to record the answers verbatim.

After administering all of the picture (word) problems, say: **"Now we are going to try something different. I am going to have you complete some more problems just like the ones you just finished, but this time the problems will be presented in word (picture) format. Try this one. Just like this word (picture) goes with this word (picture), this word (picture) goes with which of these words (pictures) down here?".**

Do not offer any coaching when presenting the second set of problems.