Alzheimer's Disease Cooperative Study

# ADAS – Cognitive Behavior

SAMPLE FORM – Page 1 of 4

| Center Name | Patient Number | Patient Initials | Examiner Initials | Examination Date |
|---|---|---|---|---|
| | P R – – | | | Month Day Year |

**1. WORD RECALL TASK**: Indicate the total number of ***correct*** responses for each trial

| Trial 1 | Trial 2 | Trial 3 |
|---|---|---|
| | | |

**2. NAMING OBJECTS AND FINGERS:** Check each object/finger named ***correctly*** or check "NONE."

NONE ☐

| | | |
|---|---|---|
| ☐ Flower | ☐ Rattle | ☐ Wallet |
| ☐ Bed | ☐ Mask | ☐ Harmonica |
| ☐ Whistle | ☐ Scissors | ☐ Stethoscope |
| ☐ Pencil | ☐ Comb | ☐ Tongs |
| ☐ Thumb | ☐ Index | ☐ Ring |
| ☐ Pinky | ☐ Middle | |

**3. COMMANDS:** Check each command performed ***correctly*** or check "NONE." NONE ☐

- ☐ Make a fist.
- ☐ Point to the ceiling, then to the floor.
- ☐ Put the pencil on top of the card, then put it back.
- ☐ Put the watch on the other side of the pencil and turn over the card.
- ☐ Tap each shoulder twice with two fingers keeping your eyes shut.

**4. CONSTRUCTIONAL PRAXIS:** Check each figure drawn ***correctly***.

- ☐ None: attempted but drew no forms correctly.
- ☐ Patient drew no forms; scribbled; wrote words.
- ☐ Circle
- ☐ Two overlapping rectangles
- ☐ Rhombus
- ☐ Cube

**5. IDEATIONAL PRAXIS:** Check each step completed ***correctly*** *or* check "NONE" NONE ☐

- ☐ Fold a letter.
- ☐ Put letter in envelope.
- ☐ Seal envelope.
- ☐ Address envelope.
- ☐ Indicate where stamp goes.

**6. ORIENTATION:** Check each item answered ***correctly*** or check "NONE." NONE ☐

| | |
|---|---|
| ☐ Full name | ☐ Day |
| ☐ Month | ☐ Season |
| ☐ Date | ☐ Place |
| ☐ Year | ☐ Time of day |

**7. WORD RECOGNITION TASK**: Scoring will be done by the A.D.C.S. Data Coordinating Center.

| Trial 1 | Trial 2 | Trial 3 |
|---|---|---|
| X | X | X |

**8. LANGUAGE:** Check level of impairment.

- ☐ **None**: patient speaks clearly and/or is understandable.
- ☐ **Very Mild**: one instance of lack of understandability.
- ☐ **Mild**: patient has difficulty < 25% of the time.
- ☐ **Moderate**: patient has difficulty 25–50% of the time.
- ☐ **Moderately Severe**: patient has difficulty more than 50% of the time.
- ☐ **Severe**: one- or two-word utterances; fluent, but empty speech; mute.

**9. COMPREHENSION OF SPOKEN LANGUAGE:** Check level of impairment

- ☐ **None**: patient understands.
- ☐ **Very Mild**: one instance of misunderstanding.
- ☐ **Mild**: 3–5 instances of misunderstanding.
- ☐ **Moderate**: requires several repetitions and rephrasing.
- ☐ **Moderately Severe**: patient only occasionally responds correctly; i.e., yes – no questions.
- ☐ **Severe**: patient rarely responds to questions appropriately; not due to poverty of speech.

**10. WORD FINDING DIFFICULTY:** Check one response.

- ☐ **None.**
- ☐ **Very Mild**: 1 or 2 instances, not clinically significant.
- ☐ **Mild**: noticeable circumlocution or synonym substitution.
- ☐ **Moderate**: loss of words without compensation on occasion.
- ☐ **Moderately Severe**: frequent loss of words without compensation.
- ☐ **Severe**: nearly total loss of content words; speech sounds empty; 1– to 2-word utterances.

**11. REMEMBERING TEST INSTRUCTIONS:** Check level of impairment.

- ☐ **None**.
- ☐ **Very Mild**: forgets once.
- ☐ **Mild**: must be reminded 2 times.
- ☐ **Moderate**: must be reminded 3–4 times.
- ☐ **Moderately Severe**: must be reminded 5–6 times
- ☐ **Severe**: must be reminded 7 or more times.

Alzheimer's Disease Cooperative Study

# ADAS – Word Recall

SAMPLE FORM – Page 2 of 4

| Center Name | Patient Number | Patient Initials | Examiner Initials | Examination Date |
|---|---|---|---|---|
| | P R – – | | | Month Day Year |

**Present Word List #2.**

**Check EACH word correctly recalled.**

| TRIAL 1 | |
|---|---|
| BOTTLE | |
| POTATO | |
| GIRL | |
| TEMPLE | |
| STAR | |
| ANIMAL | |
| FOREST | |
| LAKE | |
| CLOCK | |
| OFFICE | |
| **TOTAL** | |

| TRIAL 2 | |
|---|---|
| FOREST | |
| TEMPLE | |
| BOTTLE | |
| STAR | |
| POTATO | |
| GIRL | |
| CLOCK | |
| ANIMAL | |
| LAKE | |
| OFFICE | |
| **TOTAL** | |

| TRIAL 3 | |
|---|---|
| GIRL | |
| TEMPLE | |
| POTATO | |
| ANIMAL | |
| FOREST | |
| LAKE | |
| OFFICE | |
| CLOCK | |
| BOTTLE | |
| STAR | |
| **TOTAL** | |

Indicate total number of words correctly recalled for EACH trial on the ADAS Cognitive Behavior Form.

**12. Executive Function** (Maze):

a. ☐☐ number of errors

b. ☐☐ time at completion or second error (total seconds)

**13. Number Cancellation:**

a. ☐☐ number of targets hit (Range: 0 - 40)

b. ☐☐ number of errors

c. ☐☐ number of times to remind of task

If any item(s) 1-13 are incomplete or not done, please specify reason:

☐ Subject too cognitively impaired to complete

☐ Subject was unable to complete for physical reasons

☐ Subject refused

☐ Not Done, for reason other than above explain: _____

_____

_____

_____

_____

_____

Alzheimer's Disease Cooperative Study

# ADAS – Delayed Recall

SAMPLE FORM – Page 3 of 4

| Center Name | Patient Number | Patient Initials | Examiner Initials | Examination Date |
|---|---|---|---|---|
| | P R – – | | | Month Day Year |

**Instructions:** Say to the patient, **"NOW I WANT YOU TO TRY TO REMEMBER THE WORDS THAT I SHOWED YOU EARLIER ON PRINTED CARDS. CAN YOU TELL ME ANY OF THOSE WORDS?"**

Allow a maximum of two minutes for recall.

**check EACH word correctly recalled.**

| | |
|---|---|
| BOTTLE | |
| POTATO | |
| GIRL | |
| TEMPLE | |
| STAR | |
| ANIMAL | |
| FOREST | |
| LAKE | |
| CLOCK | |
| OFFICE | |
| **TOTAL** | |

Alzheimer's Disease Cooperative Study

# ADAS – Word Recognition

SAMPLE FORM – Page 4 of 4

| Center Name | Patient Number | Patient Initials | Examiner Initials | Examination Date |
|---|---|---|---|---|
| | P R – – | | | Month Day Year |

**Present Word List #2.**

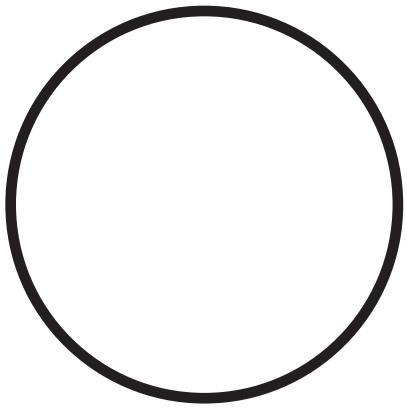
**Check subject's response for each word. Subject should respond "yes" to original words which are bolded. INCORRECT responses are shaded. Three trials of reading and recognition are given.**

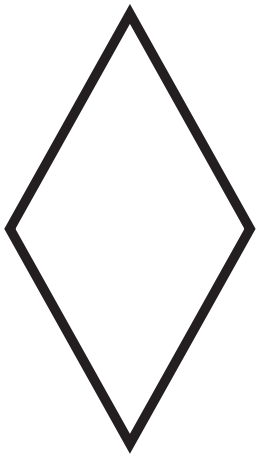
| | Yes | No |
|---|---|---|
| **COST** | | |
| NATION | | |
| **CHIMNEY** | | |
| SPARROW | | |
| **DAMAGES** | | |
| TRAFFIC | | |
| **SANDWICH** | | |
| SERVICE | | |
| SHELL | | |
| **SOLUTION** | | |
| YARD | | |
| **TUBE** | | |
| BODY | | |
| GROUND | | |
| STICK | | |
| **ENGINE** | | |
| **RICHES** | | |
| **GRAVITY** | | |
| SUMMER | | |
| WISDOM | | |
| MAN | | |
| **MEAL** | | |
| **PASSENGER** | | |
| **ACID** | | |

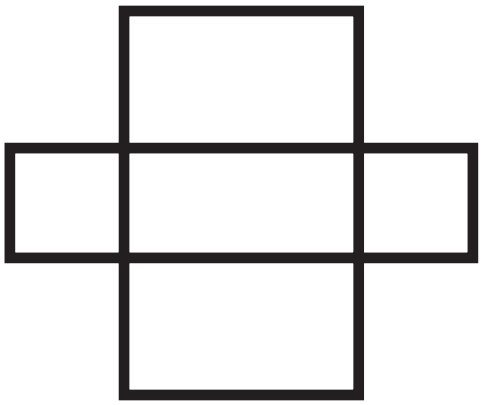
| | Yes | No |
|---|---|---|
| BATTLE | | |
| MUCH | | |
| **TUBE** | | |
| TEAM | | |
| COPY | | |
| **ENGINE** | | |
| **GRAVITY** | | |
| **COST** | | |
| JAR | | |
| DISTANCE | | |
| TRIUMPH | | |
| TEMPER | | |
| SENTENCE | | |
| FOX | | |
| **PASSENGER** | | |
| **SANDWICH** | | |
| **SOLUTION** | | |
| WHISTLE | | |
| **CHIMNEY** | | |
| UNION | | |
| **ACID** | | |
| **MEAL** | | |
| **DAMAGES** | | |
| **RICHES** | | |

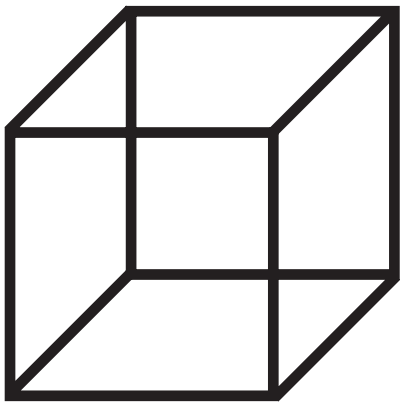
| | Yes | No |
|---|---|---|
| VISITOR | | |
| **ACID** | | |
| SPEAK | | |
| **SOLUTION** | | |
| NAME | | |
| **MEAL** | | |
| LINE | | |
| BILL | | |
| **CHIMNEY** | | |
| **ENGINE** | | |
| WEALTH | | |
| **TUBE** | | |
| IMAGE | | |
| **COST** | | |
| **SANDWICH** | | |
| **DAMAGES** | | |
| ELEPHANT | | |
| **RICHES** | | |
| **GRAVITY** | | |
| FUTURE | | |
| **PASSENGER** | | |
| STRING | | |
| BANNER | | |
| BERRY | | |

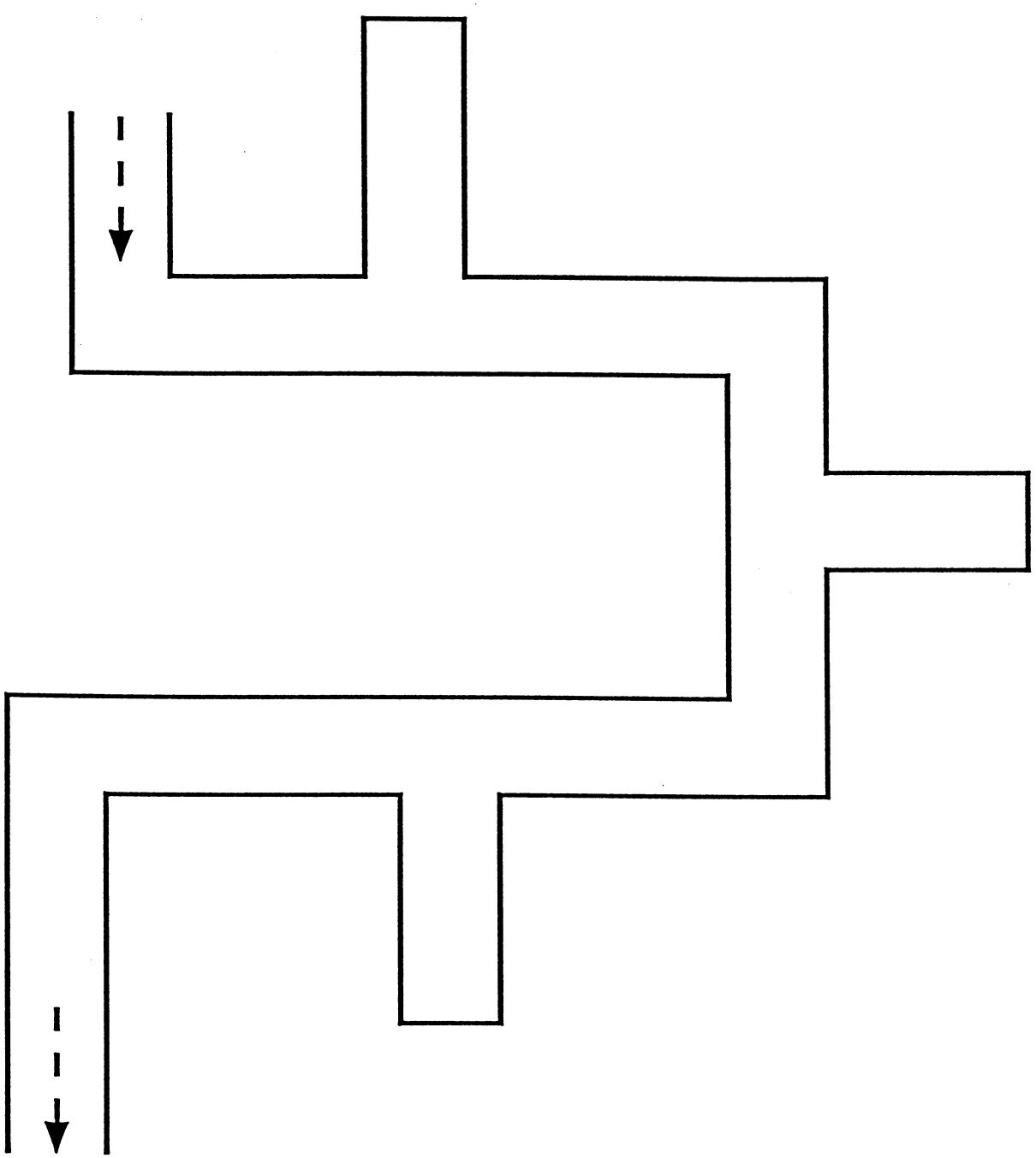
*see procedures manual for further clarification





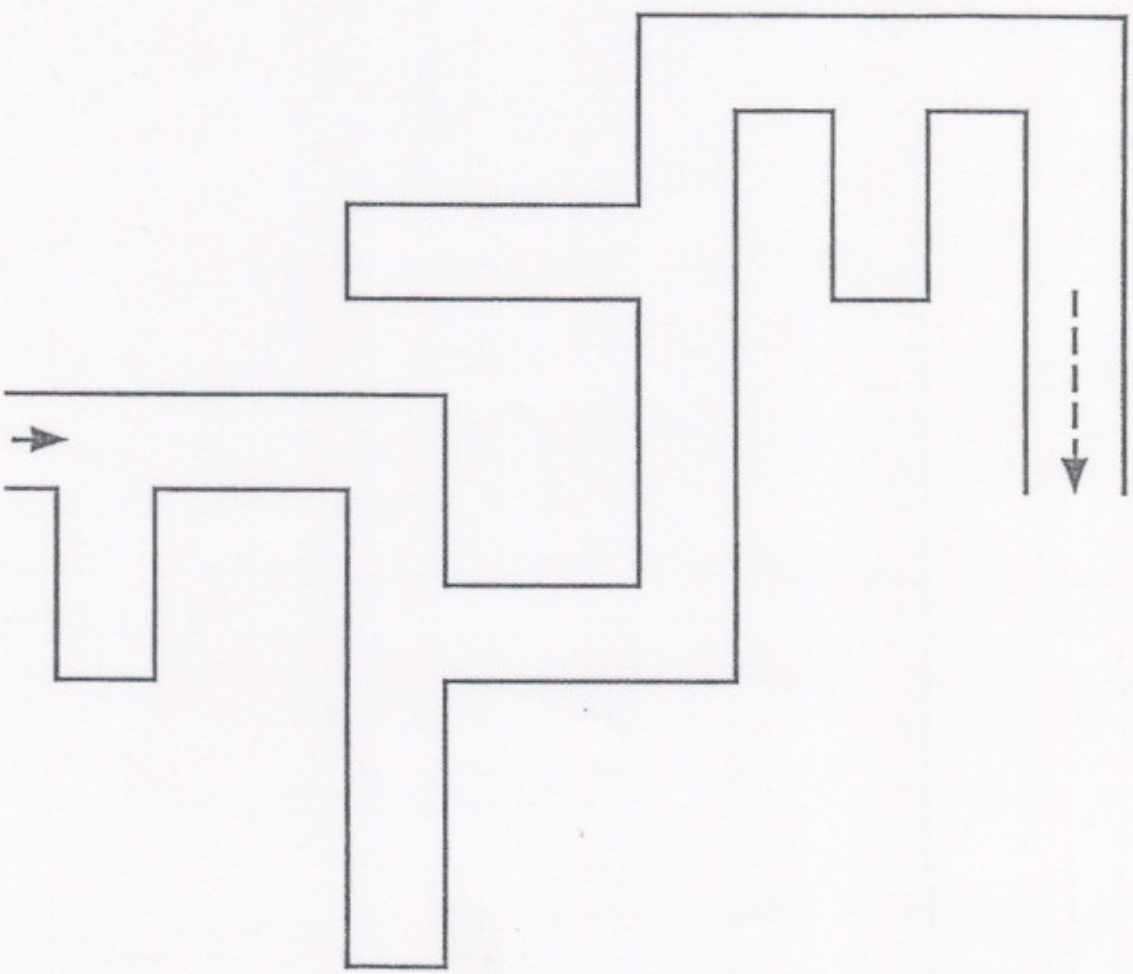






**Example**

A2



"6" and "1"

1 2 2 4 5 9 5 6 6 9 1 9 6 7 8 3 2 4 3 7 2 1 4 2 2 1 2 6 6 3

**Example for use with versions A, B, and C**

"2" and "8"

6 2 6 7 2 3 1 3 8 5 5 5 8 1 7 9 1 7 2 7 4 5 7 6 1 3 9 6 2 1

9 4 6 9 5 7 1 8 9 5 6 5 4 2 7 1 5 2 7 9 1 7 1 1 1 4 2 8 5 8

1 9 7 9 7 1 6 7 8 6 5 5 7 2 9 6 5 9 5 4 7 3 2 4 5 6 1 4 3 4

4 6 8 4 1 4 1 7 2 4 7 1 7 6 7 5 4 9 8 7 5 6 2 1 6 9 3 1 4 8

7 8 6 7 1 7 1 3 4 3 9 8 6 5 1 8 3 4 2 6 9 9 6 1 6 4 3 9 3 4

4 9 3 8 7 2 5 4 4 8 7 6 4 1 4 7 2 6 8 7 5 6 3 2 6 4 4 6 8 4

4 8 3 4 7 5 4 4 7 9 7 3 6 8 6 5 4 7 4 3 4 9 2 5 3 5 4 7 3 5

4 9 3 3 8 1 8 4 2 6 5 6 6 1 7 2 4 2 9 7 9 7 6 1 5 1 4 1 9 8

**A**

# ADAS

# Administration Manual for the Alzheimer's Disease Assessment Scale

Adapted from the Administration and Scoring Manual for the
Alzheimer's Disease Assessment Scale,
1994 Revised Edition, Richard C. Mohs, Ph.D.


Present manual modified by:
Donald Connor, Ph.D
Kimberly Schafer, MS
(3/98)

A Publication of the
Alzheimer's Disease Cooperative Study

## INTRODUCTION

The test items on the cognitive part of the **ADAS** should be given in the order indicated.

**The WORD RECALL test is given first**
and the
***WORD RECOGNITION* task is given last**
with the
**other cognitive tests given in-between.**

Separating the two word memory tasks in this way minimizes the chance that a subject will confuse the words from the two tasks.

At the start of a test session, before giving the **WORD RECALL** test, the tester should have a short conversation with the subject about neutral topics such as the weather, the subject's trip to the clinic, or what the subject had for breakfast. This conversation will help to put the subject at ease before the testing begins and will give the tester an opportunity to observe how the subject can use and understand language.

- There are three clinical ratings of language ability on the cognitive part of the **ADAS**.

The **ADAS** is not a timed test and the subject's score does not depend upon how rapidly the test is completed. The cognitive items should be given so that the session moves smoothly and quickly, but not so that the subject feels pressured to respond rapidly.

Feedback to the subject should be neutral and, usually, should not indicate whether or not the response was correct. Comments such as, **"That's fine"** or **"You're doing well"** are appropriate as long as the subject is trying. If the subject specifically asks whether or not they were correct, feedback can be given.

## TABLE OF CONTENTS

## INSTRUCTIONS for WORD RECALL TASK

**On this task, the subject is given three trials to learn a list of high-frequency, high-imagery nouns. The 10 words are printed in block letters on white cards.**

Use the appropriate word list for each visit as indicated on the study worksheet, and record the subject's responses on the study worksheet.

| | |
|---|---|
| At the start of the first trial, the tester gives instructions similar to the following: | **"I am going to show you some words printed on these white cards one at a time. Please read each word out loud and try to remember it, because later I will ask you to try to remember all of the words I have shown you. Ready, read the word and try to remember it."** |
| The examiner can prompt with: | **"Read it out loud and try to remember it"** as necessary. |

If the subject cannot read the word or is slow, the examiner can say the word out loud and have the subject repeat it. Note this and continue with this procedure at each testing. In some cases, the examiner may have to say all of the words and have the subject repeat them. Regardless, make sure the subject looks at each word while repeating it.

| | |
|---|---|
| After the presentation, the tester asks the subject to try to recall as many of the words as possible by saying: | **"Good, now tell me all the words you remember that were on the list."** |

Two more learning and recall trials follow.

| | |
|---|---|
| For trials 2 and 3, say to the subject: | **"Now I'm going to show you that same list again. Read each word out loud and try to remember it."** |

Encouragement can be given if the subject is nervous or giving up.

**SCORING:**

**The subject's score is the mean number of words *not* recalled on three trials (maximum score = 10)**

**Enter the subject's score on the study worksheet**

## INSTRUCTIONS for NAMING TASK

**For this task, the subject is asked to name the 12 randomly presented real objects, with**
**high** (Flower, Bed, Whistle, Pencil),
**medium** (Rattle, Mask, Scissors, Comb), and
**low** (Wallet, Harmonica, Stethoscope, Tongs) **frequency values.**

**The subject is also asked to name the fingers on his/her dominant hand.**

- Use the study worksheet to record the subject's responses.
- Objects should be presented in random order. Do not allow the subject to touch the objects.

| | |
|---|---|
| Give the subject instructions similar to the following: | **"Now I am going to show you some objects. I want you to tell me what their names are. What is this called?"** (present object) |
| Continue to present objects in random order. The first question about each object should be: | **"What is this called?"** or **"What is the name of this thing?"** |
| If the subject responds with the object's function say: | **"Yes, that's what it does, but what is its name?"** |

- If the subject does not respond, the examiner should give the clue for that item provided below. If the subject still does not respond or makes an error, go on to the next object.

| ITEM | CLUES |
|---|---|
| **Flower** | grows in a garden |
| **Bed** | used for sleeping in |
| **Whistle** | makes a sound when you blow on it |
| **Pencil** | used for writing |
| **Rattle** | a baby's toy |
| **Mask** | hides your face |
| **Scissors** | cuts paper |
| **Comb** | used on hair |
| **Wallet** | holds your money |
| **Harmonica** | a musical instrument |
| **Stethoscope** | doctor uses it to listen to your heart |
| **Tongs** | picks up food |

## INSTRUCTIONS for NAMING TASK (Cont'd.)

➲ The subject is also asked to name the fingers of his/her dominant hand (*e.g.,* thumb, index [pointer/forefinger], middle, ring finger, and pinky).

| | |
|---|---|
| Give the subject instructions similar to the following: | **"Now I am going to point to a part of your hand and I want you to tell me what it's called. What is this?"** |
| For the 4 fingers, if a query is necessary, say: | **"What is another name for this finger?"** |

| ITEM |
|---|
| **Thumb** |
| **Index**/forefinger/pointer |
| **Middle** |
| **Ring** |
| **Pinky** |

The hardest part of scoring the naming task is determination of the range of correct responses based on the subject's cultural and geographical background. A response other than the name given on the response form should be scored as correct if it is a name that would be used by a non-demented person with the same cultural background as the subject.

**FOR EXAMPLE:** the **Mask** might be called a **"false face"** in some parts of the U.S.; the **Wallet** might be called a **"billfold"** or the **Harmonica** might be called a **"mouth organ"**.

➲ Descriptions of the object, semantic or phonemic paraphasias should not be scored as correct.

**EXAMPLES OF INCORRECT RESPONSES ARE:** **"listening thing"** for **Stethoscope,** **"cutter"** for **Scissors,** and **"prongs"** for **Tongs.**

| SCORING | |
|---|---|
| **0 = 0-2** | items (objects and fingers) named incorrectly |
| **1 = 3-5** | items (objects and fingers) named incorrectly |
| **2 = 6-8** | items (objects and fingers) named incorrectly |
| **3 = 9-11** | items (objects and fingers) named incorrectly |
| **4 = 12-14** | items (objects and fingers) named incorrectly |
| **5 = 15-17** | items (objects and fingers) named incorrectly |
| **Enter the subject's score on the study worksheet.** | |

## INSTRUCTIONS for COMMANDS

**This task is designed to assess receptive speech. The subject is asked to carry out 5 separate commands with 1 to 5 steps per command.**

- Each command should be read once. If the subject does not respond or makes an error, the tester should give the ***ENTIRE*** command one more time.
- All commands should be given to every subject.
- If the subject demonstrates hearing or attentional difficulties, orient them by saying, **"Ready?"** or **"Now I want you to..."** prior to giving the command. Do ***NOT*** give the command more than twice.
- There should be no other materials near the pencil, watch and card (pens, paper, etc.)
- Each underlined element represents a single step.
- Each command is scored as a whole (no partial credit). All components must be correct for the response to be scored as correct.
- Use the study worksheet to record the subject's responses.

| | |
|---|---|
| Give the subject instructions similar to the following: | **"Now I am going to ask you to do a few things. First, ...**<br>**"Make a *FIST*."** (**"Relax it"** if needed)<br>**"Point to the *CEILING* and then to the *FLOOR*."** |
| Line up a Pencil, Watch, and Card on the table. Say: | **"Put the *PENCIL ON TOP OF THE CARD* and then *PUT IT BACK*."**<br>**"Put the *WATCH* on the *OTHER SIDE OF THE PENCIL* and then *TURN OVER THE CARD*."** |
| Remove the Pencil, Watch, and Card from the table. Say: | ***"TAP EACH SHOULDER TWICE* with *TWO FINGERS* keeping your *EYES SHUT*."** |

| SCORING | |
|---|---|
| 0 = | All commands correct |
| 1 = | 1 command incorrect, 4 commands correct |
| 2 = | 2 commands incorrect, 3 commands correct |
| 3 = | 3 commands incorrect, 2 commnads correct |
| 4 = | 4 commands incorrect, 1 command correct |
| 5 = | All 5 commands correct |

**Enter the subject's score on the study worksheet**

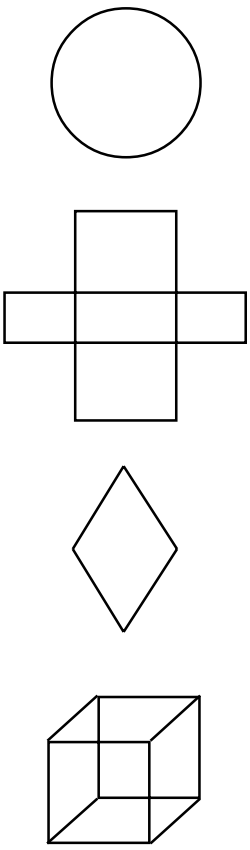
## INSTRUCTIONS for CONSTRUCTIONAL PRAXIS

**This test assesses the subject's ability to copy 4 geometric forms ranging from a very simple one (circle) to a fairly difficult one (cube).**

- Use the study worksheet to score the subject's responses.
- The forms provided should be presented one at a time.
- The tester should give the subject a lead pencil with an eraser along with the drawing.

| | |
|---|---|
| The instructions to the subject should be similar to the following: | **"On this piece of paper is a shape. Try to draw another one that looks just like this, somewhere on the page."** (Examiner may point to shape) |
| If the subject's response is quick or sloppy, prompt with: | **"Take your time and try to draw it just like this one."** |

- The subject should be allowed **two attempts** for each shape. Allow a second attempt only if the subject asks or indicates a problem with their drawing. The subject may erase if they need to. If the subject draws on top of the printed design, count this as one attempt and indicate that they should try on an empty part of the page. If the subjects says the reproduction is poor, query if the subject wants another try. When two attempts are made, ask the subject to indicate which one is the best, and then score that attempt.
- If the subject cannot reproduce the figure in two attempts, the tester should go on to the next item.
- A drawing should be scored as correct if the subject has reproduced all of the essential features of the original. Changes in size do not count as errors. Small gaps between lines do not indicate an error, as long as the shape has been reproduced.



## INSTRUCTIONS for CONSTRUCTIONAL PRAXIS (cont'd)

The forms should be presented in the following order:

| |
|---|
| **Circle** |
| **Two Overlapping Rectangles** |
| **Diamond (Rhombus)** |
| **Cube** |

### SCORING GUIDELINES:

**Circle:** A closed curved figure

**Two Overlapping Rectangles:** Forms must be four-sided, and overlap must be similar to presented form. Changes in size are not scored.

**Diamond (Rhombus):** Figure must be four-sided, oriented so that the points are at the top and bottom, and the sides are approximately equal length (e.g., longest side is not ≥ 1.5 times the length of the shortest side).

**Cube:** The form is 3-dimensional, with front face in the correct orientation, internal lines drawn correctly between corners. Opposite sides of faces should be approximately parallel.

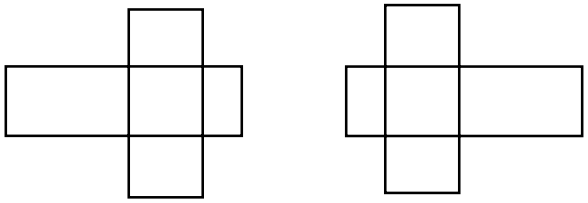
| SCORING | |
|---|---|
| 0 = | All 4 drawings correct |
| 1 = | 1 form drawn incorrectly |
| 2 = | 2 forms drawn incorrectly |
| 3 = | 3 forms drawn incorrectly |
| 4 = | 4 forms drawn incorrectly |
| 5 = | No figures drawn, scribbles; parts of forms; words instead of forms |

**Enter the subject's score on the study worksheet**

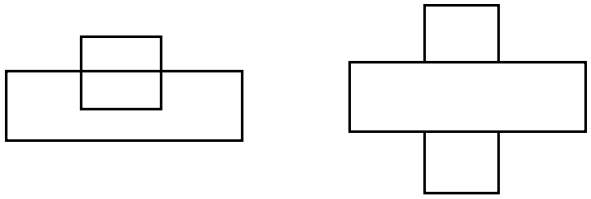
### Examples of Correct and Incorrect Drawings:

#### Overlapping Rectangles

Correct

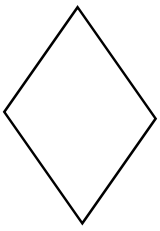


Incorrect



#### Diamond

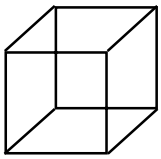
Correct



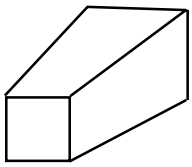
Incorrect

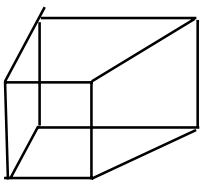
#### Cube

Correct



Incorrect





## INSTRUCTIONS for IDEATIONAL PRAXIS

**This task is designed to determine whether the subject can perform a familiar but complex sequence sequence of actions.**

- Use the study worksheet to record the subject's responses.
- A long envelope, an 8.5" x 11" sheet of paper and a pencil are placed in front of the subject.

Give the subject instructions similar to the following:

**"I want you to pretend you have written yourself a letter. Take this piece of paper, fold it so that it will fit into the envelope, and then put it into the envelope. Then, seal the envelope, address the envelope to yourself, and show me where the stamp goes."**

- There are 5 components to this task and each one is underlined in the instruction.
- If the subject forgets part of the task, or is having difficulty, the tester should repeat the instruction for the component of the task where the subject is having difficulty.

**FOR EXAMPLE:** If the subject stops after folding the paper and putting it in the envelope, the tester should give **one reminder** on the next component:, **"Now seal the envelope."** If the subject cannot do this part, move on and give one reminder on the next component:**"Now address the envelope to yourself."**

After the first complete instruction only one additional reminder should be given for each component.

- Impairment on this item should reflect dysfunction in executing an overlearned task only and not recall difficulty.
- Any address which would enable a postal worker to deliver the envelope is counted as correct, even though it might not contain the subject's current address. The address should contain: name, street, city, and state. Zip code is not required.
- Have the subject indicate where the stamp goes by placing an "X" on the envelope.

| SCORING | |
|---|---|
| 0 = | All components performed correctly |
| 1 = | Failure to perform 1 component |
| 2 = | Failure to perform 2 components |
| 3 = | Failure to perform 3 components |
| 4 = | Failure to perform 4 components |
| 5 = | Failure to perform 5 components |

**Enter the subject's score on the study worksheet**

## INSTRUCTIONS for ORIENTATION

**This task is designed to determine how well oriented the subject is with regard to time and place.**

- Use the study worksheet to record the subject's responses.
- The components of orientation are:

| |
|---|
| **Person** |
| **Day of the Week** |
| **Date, Month, Year** |
| **Season** |
| **Time of Day** |
| **Place** |

- The tester should ask the subject for each of these pieces of information one at a time.
- Make sure no watches, clocks, calendars, etc. are visible to the subject.
- One restatement of question is allowed (*e.g.*, if subject confuses day and date).
- Acceptable range for answers include:

| | |
|---|---|
| **Date:** | + one day |
| **Time:** | + one hour |
| **Place:** | Partial name acceptable (*e.g.*, name of hospital, clinic or professional building) |
| **Season:** | Within one week prior to onset or within two weeks of termination |

- **Month, Year, Day of the Week**, and the **subject's first and last name** must be **exact.**

**SCORING:** One point is given for each incorrect response (maximum = 8)

- Enter the subject's score on the study worksheet.

## INSTRUCTIONS for WORD RECOGNITION

**On this task the subject is given one trial to learn a list of 12 words.**

- Use the appropriate word list as indicated on the study worksheet, and record the subject's responses on the study worksheet.
- The learning part of this trial is similar to the learning part of the WORD RECALL TEST since the subject is asked to read each word aloud and try to remember it.
- For the one test trial, the 12 studied words are mixed with 12 new words matched to the studied words for frequency and imagery and the subject is asked to decide for each word whether or not it was one of the studied words.

| | |
|---|---|
| At the start of the Learning Trial, give the subject instructions similar to the following: | **"I am going to show you some words printed on these white cards. I want you to read each word out loud and try to remember it."** |

Some of the words on the WORD RECOGNITION TASK may not be familiar to the subject and the subject may have difficulty reading them. If the subject cannot read a word, the tester should say the word out loud. However, it is important for the subject to actually look at each word and try to read it.

| | |
|---|---|
| At the end of the learning portion of a trial the tester should say something to the subject similar to the following: | **"Now I'm going to show you another set of words. Some of the words were on the list I just showed you and others are new. For each word I want you to tell me whether it is one of the words I just showed you."** |
| The tester shows the first word and says either:<br>or: | **"Is this one of the words I showed you before, yes or no?**<br>**"Did I show you this word before?"** |
| The same instruction is given before the second test word. For the remaining test words the tester should say: | **"How about this one?"** |

## INSTRUCTIONS for WORD RECOGNITION (cont'd)

- If the subject does not remember the task (*e.g.*, reads the word rather than responding **"Yes"** or **"No"**) then the tester should repeat or rephrase the entire question and make a note in the appropriate column on the worksheet that the subject had to be reminded of the task instructions. Likewise, if the subject appears to have fallen into a response set (*i.e.*, saying **"Yes"** to every word or saying **"No"** to every word), then the test instructions should be repeated.

## INSTRUCTIONS for REMEMBERING TEST INSTRUCTIONS

**This item evaluates the subject's ability to remember the requirements of the WORD RECOGNITION TASK.**

| | |
|---|---|
| On each recognition trial, the subject is asked prior to presentation of the first two words: | **"Did I show you this word before, or is this a new word?"** |
| For the third word, the subject is asked: | **"How about this one?"** |

- If the subject responds accurately, *i.e.*, **"Yes"** or **"No"**, then memory for the instructions is accurate.
- If the subject fails to respond, this signifies that the instructions have been forgotten and the instruction is repeated.
- The procedure used for the third word is repeated for words 4-24. Each instance of memory failure for the test instructions is noted.

| SCORING | |
|---|---|
| 0 = | Subject never needs extra reminders of instructions |
| 1 = | Very mild – forgets once |
| 2 = | Mild – must be reminded 2 times |
| 3 = | Moderate – must be reminded 3 or 4 times |
| 4 = | Moderately severe – must be reminded 5 or 6 times |
| 5 = | Severe – must be reminded 7 or more times |
| **Enter the subject's score on the study worksheet** | |

## INSTRUCTIONS for SPOKEN LANGUAGE ABILITY

**This item is a global rating of the quality of speech, *i.e.*, clarity, difficulty in making oneself understood.**

- In rating this item the tester should consider all of the speech produced by the subject during the test session.
- Quantity of speech and word finding difficulty are not rated on this item.
- It should be noted that the higher scores (4-5) on this item are reserved for subjects whose expressive language abilities are impaired to such an extent that they seldom communicate without difficulty.

| SCORING | |
|---|---|
| 0 = | No instances when it is difficult to understand the subject |
| 1 = | Very mild – one instance of lack of understandability |
| 2 = | Mild – subject has difficulty less than 25% of the time |
| 3 = | Moderate – subject has difficulty 25-50% of the time |
| 4 = | Moderately severe – subject has difficulty 50% of the time |
| 5 = | Severe – one or two word utterance; fluent, but empty speech; mute |

**Enter the subject's score on the study worksheet**

## INSTRUCTIONS for WORD-FINDING DIFFICULTY AND COMPREHENSION

### Word-Finding Difficulty in Spontaneous Speech

Along with Spoken Language Ability, this item rates impairment in expressive speech, but it rates ***only word finding difficulty***, whereas Spoken Language Ability is a more global rating of the extent to which the subject can communicate verbally.

- To rate this item, the tester must determine whether the subject has difficulty in finding the desired word in spontaneous speech. The problem may be overcome by circumlocution, *i.e.*, giving explanatory phrases or nearly satisfactory synonyms.
- Do not include finger and object naming in this rating.

| SCORING | |
|---|---|
| 0 = | No evidence of word finding difficulty in spontaneous speech |
| 1 = | Very mild – 1 or 2 instances, not clinically significant |
| 2 = | Mild – noticeable circumlocution or synonym substitution |
| 3 = | Moderate – loss of words without comprehension on occasion |
| 4 = | Moderately severe – frequent loss of words without comprehension |
| 5 = | Severe – near total loss of content of words; speech sounds empty; 1 – 2 word utterances |
| **Enter the subject's score on the study worksheet** | |

### Comprehension — This item rates the subject's ability to understand speech

- To rate this item, the tester should consider how well the subject was able to understand the tester's speech during the opening discussion and during the test session
- Do not include responses to commands

| SCORING | |
|---|---|
| 0 = | No evidence of poor comprehension |
| 1 = | Very mild – 1 or 2 instances of misunderstanding |
| 2 = | Mild – 3-5 instances of misunderstanding |
| 3 = | Moderate – requires several repetitions and rephrasing |
| 4 = | Moderately severe – subject only occasionally responds correctly, *i.e.*, yes/no questions |
| 5 = | Severe – subject rarely responds to questions appropriately, not due to poverty of speech |
| **Enter the subject's score on the study worksheet** | |

*Alzheimer Disease and Associated Disorders*
Vol. 11, Suppl. 2, S13–S21


# Development of Cognitive Instruments for Use in Clinical Trials of Antidementia Drugs: Additions to the Alzheimer's Disease Assessment Scale That Broaden Its Scope

Richard C. Mohs, *David Knopman, †Ronald C. Petersen, ‡Steven H. Ferris, §Chris Ernesto, §Michael Grundman, ‖Mary Sano, ¶Linas Bieliauskas, **David Geldmacher, ††Chris Clark, §Leon J. Thal, and the Alzheimer's Disease Cooperative Study

*Department of Psychiatry, Mount Sinai Medical School, Bronx, New York; *Department of Neurology, University of Minnesota, Minneapolis, Minnesota; †Department of Neurology, Mayo Clinic, Rochester, Minnesota; ‡Department of Psychiatry, New York University, New York, New York; §Department of Neurosciences, University of California, San Diego, California; ‖Sergievsky Center, Columbia University, New York, New York; ¶Department of Neurology, University of Michigan, Ann Arbor, Michigan; **Department of Neurology, Case Western Reserve University, Cleveland, Ohio; and ††Department of Neurology, University of Pennsylvania, Philadelphia, Pennsylvania, U.S.A.*

**Summary:** The cognitive assessment protocol of the Alzheimer's Disease Cooperative Study (ADCS) was designed to evaluate the reliability and validity of cognitive assessment measures that might be valuable additions to the Alzheimer's Disease Assessment Scale (ADAS) or other concise batteries used in antidementia drug trials. As part of an overall ADCS protocol to develop new instruments to be used in trials of treatments for Alzheimer's disease (AD), patients with mild to moderate AD and cognitively normal elderly were administered a battery of five tests at least three times over 1 year. The tests included word list learning with delayed free recall, a recognition memory test for faces, a series of letter and digit cancellation tests to measure concentration, tests of praxis, and a series of maze completion tasks designed to assess planning and executive function. A version of the digit cancellation task was reliable and sensitive to a broad range of dementia severity so that it could provide a useful addition to the present version of the ADAS. Performance on the word learning task with delayed recall and a subset of the mazes task were impaired even in mild AD, so these tasks may be useful in trials involving mild or at-risk subjects. Performances on the facial recognition task and on the praxis tasks were not related to dementia severity, so these tasks would not be useful to evaluate treatments. Therefore, the major outcome of this investigation was the identification of some potential addtions to the present ADAS that extend both the cognitive domains and the range of symptom severity covered. **Key Words:** Assessment scales—Alzheimer's disease—Cognitive impairment.

Although patients with Alzheimer's disease (AD) may have a variety of clinical symptoms, a progressive impairment in memory and other cognitive functions is their most prominent characteristic. Recent efforts to develop more effective treatments for AD have been concerned primarily with treatments that might improve cognitive function or, possibly, slow the rate of cognitive decline. As knowledge about the pathophysiology of AD has improved and as the number of potential new drug treatments has increased, the assessment of the cognitive effects of drug treatments has also evolved. Early clinical trials used measures borrowed from the diagnostic neu-

Address reprint requests to Dr. R. C. Mohs at Psychiatry Service (116A), Mount Sinai School of Medicine, Veterans Affairs Medical Center, 130 W. Kingsbridge Road, Bronx, NY 10468, U.S.A.

ropsychology laboratory (Yesavage et al., 1979) whereas more recent studies have used one or two comprehensive instruments specifically designed to grade the severity of dementia such as the Mini-Mental State Examination (MMSE) (Folstein et al., 1975) and the Alzheimer's Disease Assessment Scale (ADAS) (Rosen et al., 1984). The advantages of a single instrument are several. They include efficiency of administration, restricted time demands for patients, conservation of statistical power, and simplicity of interpretation.

The ADAS was developed with these features in mind, specifically for longitudinal assessment of AD patients. It has been widely used and well received. However, the ADAS does not contain certain cognitive elements that, in retrospect, might be important. In particular, the ADAS does not include any tests designed to assess attention and concentration, nor does it include any specific assessment of planning or executive functions, both of which may be impaired. Although the ADAS does assess memory in some detail, recent work (Welsh et al. 1991; Petersen et al., 1995) indicates that aspects of verbal memory such as delayed recall, which were not included in the ADAS, are important for measuring memory impairment early in the course of dementia. Memory for nonverbal material probably has a different biologic substrate than does verbal memory, and the ADAS does not include any tests of nonverbal memory. Finally, praxis can be assessed in a variety of different ways (Goodglass and Kaplan, 1972), most of which are not included in the ADAS.

The purpose of the present investigation was to investigate the validity and reliability of cognitive measures that assess those aspects of cognitive function not represented in the ADAS. The overall design of the ADCS instrument protocol enabled us to obtain the following information about each of these proposed new tests:

1. What is the 1-month test–retest reliability of the test?
2. To what extent is test performance affected by learning, as might occur in a treatment trial when patients are tested repeatedly?
3. Are there confounding effects of age and education on change scores?
4. How useful is the test across the spectrum of disease severity?
5. How sensitive is the test to change over 12 months, and how variable is that change?

## METHODS

### Overview

A detailed description of the design of the ADCS Instrument Study Protocol and of the study participants is given in the article by Ferris et al. (this issue). Subjects who participated in the cognitive instrument protocol included 64 normal elderly control (NEC) subjects, 50 mild AD patients with baseline MMSE scores ≥21 (group AD I), 47 moderate AD patients with baseline MMSE scores of 16–20 (group AD II), and 46 moderately severe AD patients with baseline MMSE scores of 10–15 (group AD III). Half of the patients in each severity group were tested at baseline, and at 1, 2, 6, and 12 months, and the other half were tested at baseline, 1, 6, and 12 months. Whenever possible, different forms of the tests were given at each test session.

The standard instruments used to characterize the severity of dementia in all subjects are described in the article by Ferris et al. (this issue). They include the Clinical Dementia Rating (CDR; Hughes et al., 1982), the Global Deterioration Scale (GDS; Reisberg et al., 1982), and the MMSE (Folstein et al., 1975).

### New Instruments

The novel cognitive instruments tested in this study were administered in the order in which they are described below. Before the study was initiated, a detailed Administration and Scoring Manual was developed to help psychometricians at the ADCS sites administer and score the tests in a uniform manner. In addition, a videotape showing a standard administration of the test battery was made and distributed to all sites. Questions concerning administration and scoring that arose during the conduct of the study were answered promptly by the ADCS Cognitive Instrument Committee and updates to the Manual were distributed to all sites.

#### *Word List Learning*

For this task, the subject was given four trials to learn a list of 10 concrete nouns. The methods were similar to those used both by the Consortium to Establish a Registry for Alzheimer's Disease (CERAD; Morris et al., 1989) and the ADAS (Rosen et al., 1984). Each of the 10 words was printed in large letters on a card. Subjects were shown the cards one at a time for 2 s each, and the subject was instructed to read each word aloud. If the subject could not read or misread a word, the examiner read it aloud. After the list was read, the subject was asked to recall the words aloud. Five different word lists, equated for mean frequency of occurrence (Thorndyke and Lorge, 1944) and for imagery (Pavio et al., 1968), were used for the five test sessions.

#### *Praxis*

The items used for tests of praxis were drawn from the Boston Diagnostic Aphasia Examination (Goodglass and

Kaplan, 1972). The subjects were told that they would be asked to follow some instructions. They were then asked to perform two facial commands (e.g., sniff a flower), two upper limb commands (e.g., wave goodbye), two instrumental commands (e.g., stir coffee with a spoon), and two whole-body commands (e.g., march like a soldier). Performance was scored from 0 to 3 according to the following criteria: 3 points, good performance on command; 2, approximate performance on command or good performance on imitation; 1, approximate performance on imitation or uses body part as object (e.g., uses hand as the instrument); and 0, cannot perform task. Four different forms of this task were used and subjects received the same form at baseline and 12 months. The subject's total score ranged from 0 to 24.

### *Delayed Verbal Recall*

For this task, the subject was asked to recall as many words as possible from the previously studied list. This task is similar to the delayed recall procedure used in the CERAD battery (Welsh et al., 1991). The subject's score was from 0 to 10.

### *Facial Recognition Memory*

This task used stimuli and procedures adapted from Warrington and James (1967). During the learning phase, 12 black-and-white pictures showing only a person's face were presented at a rate of one picture every 3 s. To ensure that the subjects attended to the picture they were asked to judge whether the faces were pleasant or not. Immediately after the 12 faces had been presented, 12 pairs of faces were then presented, and for each, the subject was asked to point to the face that had been shown previously. Four alternative forms were constructed from the total pool of stimuli developed by Warrington and James (1967), so that form 1 was used at baseline and then again at month 12 for all subjects. The subject's score was the number of correct responses out of 12.

### *Cancellation Test*

This task was adapted from paradigms used in cognitive (Neisser, 1964) and clinical neuropsychology (Lezak, 1982) that are designed to assess visual attention and concentration. In this task, the subject was presented with a page on which there were eight rows on letters or digits and the subject's task was to cross off as many target letters or digits as possible in 60 s. Because the difficulty of this task depends on the target, subjects were given six different versions in each session and for each the target was printed at the top of the page. A practice task was given to the subjects before beginning the task. Subjects were told to work as quickly as possible and were told not to erase if they crossed off an incorrect letter or number. The tester was allowed to remind subjects of the instructions if they made three errors in a row or forgot the instructions during the task. The six versions of the task were: (a) single letter (e.g., "P") mixed with other letters; (b) pair of two letters (e.g., "O E") mixed with other letters; (c) either of two letters (e.g., "L" or "T") mixed with other letters; (d) set of three letters (e.g., "W K M") mixed with other letters; (e) large letters (e.g., "E") mixed in with letters printed in normal size type; and (f) either of two numbers (e.g., "2" or "8") mixed in with other numbers.

There were 13–40 targets on each page depending on the test version (see Table 3). The subject's score was the number of target items correctly crossed off in 60 s minus the number of incorrectly crossed off items, and minus the number of reminders given. Four alternate forms of each task were available, so that form 1 was used at baseline and then again at month 12.

### *Maze Test*

Paper and pencil mazes and instructions were drawn from the battery of Christensen et al. (1991). Seven mazes of increasing complexity were presented in each session. Complexity was increased by increasing the number of turns that had to be made to draw a line from start to finish without crossing a maze boundary. Subjects were instructed to find their way through each maze with a pencil, without hitting a dead end, and were told that they could pause to make a decision. They were told to proceed as quickly as possible because the task was being timed. An example was given for practice at the beginning of the task and again just before the fourth maze. In the examples, the subjects were shown the entrance and exits to the practice mazes and were told not to lift their pencil from the paper. After hitting a dead end for the first time, the examiners would bring the subject back to the point of the incorrect decision and suggest that the subject try another direction. After hitting a second dead end, the examiner would go on to the next maze. After two consecutive mazes with two dead ends, the task was halted. If the subject did not complete a maze in 240 s the examiner went on to the next maze. The number of mazes completed and the times to completion were recorded. Four alternate forms of this test were available, so that form 1 was used at baseline and then again at month 12.

## RESULTS

### Overview

For each of the measures, we performed analyses designed to address each of the issues raised in the introduction. To assess reliability, both Pearson and Spearman correlation coefficients were calculated but, because the two measures were very close in every case, only results from the Pearson calculations are reported. Similarly, the effect of confounds due to age and education were also estimated by Pearson correlations. The effect of learning on each measure was examined by looking at change over 1 month, a time period during which little meaningful clinical change would usually occur. A more detailed analysis of the effects of repeated exposure comparing the two testing schedules is beyond the scope of this first report. The ability of each test and measure to assess the full range of dementia severity was examined by looking at the extent to which scores differed across all severity groups and the extent to which there were ceiling or floor effects. Ability to measure longitudinal change was assessed by looking at 12-month change. For some measures a difference effect size measure ($d$), defined as the mean change divided by the standard deviation of the change, was used to provide an index of the relative sensitivity of that measure to longitudinal change. The value of this measure is that it takes mean change scores, which are originally expressed in different units (e.g., seconds, errors, number of words recalled), and expresses each in terms of standard deviation units, thus enabling us to compare across measures. As for any longitudinal study, data are sometimes not available for follow-up visits owing to missed visits, untestability, or errors in data recording. In each of the data tables discussed below we have listed, along with mean performance data, the number of patients in each group for whom valid data were available at baseline and at 12 months. The overall rate of missing data was low but increased with severity of dementia and duration of follow-up.

### Word List Learning

Results for this task are presented in Table 1. The 1-month retest reliability was $r = 0.79$, 0.86, 0.84, and 0.86 for trials 1–4, respectively. For the sum of trials 1–4 the Pearson $r$ was 0.92. There clearly was no learning effect at 1 month, because most scores were slightly lower than at baseline. There were no significant correlations between age or education and any of the 12-month change score values for the four learning trials.

Baseline performance was different for all four severity groups indicating that, at least for the sum score, this task discriminates a broad range of dementia severity. Even among the NEC group there was relatively little evidence of a ceiling effect, except for trials 3 and 4. For patients in the most severe group (AD III), performance after 12 months was very poor, with 37% of patients giving no response and an additional 30% recalling a total of four or fewer words. Normal controls usually improved their performance slightly at the 1-year follow-up, whereas patients in all three AD groups showed worse performance on all trials. Because some word list learning paradigms use three trials (Morris et al., 1989) and others use more (Petersen et al., 1995), we compared the effect sizes ($d$) for the sum of three and four trials for AD I patients, in whom there were no ceiling or floor effects. For AD I patients the $d$ for the sum of trials 1 to 3 was 0.72 and $d$ for the sum of trials 1 to 4 was 0.73, suggesting that inclusion of the fourth trial did not add to the ability of the task to detect change.

### Delayed Recall

Retest reliability was excellent for the delayed recall score ($r = 0.93$), but the (trial 5 − trial 4) difference score was less reliable ($r = 0.61$). As indicated by the data in Table 1, there appeared to be no learning effect, because scores after 1 month were slightly lower than baseline for all groups. There was no significant correlation of age or education with the 12-month change score.

As has been found in previous studies (Welsh et al., 1991) the delayed recall measure was very good for discriminating the NEC group from AD I patients. However, because of floor effects this measure did not discriminate among the AD groups. Even among AD I patients, 20% recalled zero words at baseline and 45.5% recalled none at 12 months. Over half of the patients in AD groups II and III recalled zero words at baseline and, at 12 months, 71% of AD II patients and 83% of AD III patients recalled zero words. As Table 1 indicates, there was relatively little change over 12 months in any of the AD groups because of low baseline performance.

### Face Recognition Memory

The data for this task are presented in Table 2. The 1-month retest reliability for this task was poor ($r = 0.48$) but, as Table 2 indicates, there was no learning or other change in average performance over 1 month. Age was not associated with the 12-month change score ($r = -0.06$; $p = 0.43$), and there was no association of change with education ($r = 0.08$; $p = 0.29$). As Table 2 indicates, AD groups II and III performed very close to

**TABLE 1.** *Results of the word list learning and recall task*

| Group | Baseline: 1-month follow-up subjects | 1-month change score | Baseline: all subjects (*n*) | 12-month score for all subjects (*n*) | 12-month change (SD) |
|---|---|---|---|---|---|
| Trial 1 (0–10) | | | | | |
| NEC | 6.21 | −0.59 | 6.22 (64) | 6.35 (62) | 0.10 (1.77) |
| AD I | 2.92 | −0.24 | 3.20 (50) | 2.27 (44) | −0.93 (1.81) |
| AD II | 2.26 | −0.38 | 2.36 (47) | 1.48 (40) | −0.88 (1.13) |
| AD III | 1.27 | −0.11 | 1.29 (45) | .62 (29) | −0.66 (1.15) |
| Trial 2 (0–10) | | | | | |
| NEC | 7.59 | −0.13 | 7.70 (64) | 8.23 (62) | 0.50 (1.72) |
| AD I | 4.40 | −0.48 | 4.42 (50) | 3.43 (44) | −0.98 (2.01) |
| AD II | 3.66 | −0.58 | 3.49 (47) | 2.30 (40) | −1.19 (1.26) |
| AD III | 2.32 | −0.58 | 2.40 (45) | 1.33 (29) | −1.07 (1.18) |
| Trial 3 (0–10) | | | | | |
| NEC | 8.09 | −0.53 | 8.17 (64) | 8.52 (62) | 0.32 (1.11) |
| AD I | 5.24 | −0.88 | 5.12 (50) | 4.02 (44) | −1.02 (1.53) |
| AD II | 4.25 | −1.08 | 4.11 (47) | 2.75 (40) | −1.36 (1.76) |
| AD III | 2.42 | −0.05 | 2.64 (45) | 1.55 (29) | −1.04 (1.17) |
| Trial 4 (0–10) | | | | | |
| NEC | 8.40 | −0.09 | 8.67 (64) | 8.65 (62) | −0.05 (1.22) |
| AD I | 5.20 | −0.24 | 5.32 (50) | 3.89 (44) | −1.43 (2.26) |
| AD II | 3.95 | −0.58 | 4.11 (47) | 2.65 (40) | −1.46 (1.55) |
| AD III | 2.78 | 0.11 | 2.69 (45) | 1.52 (29) | −1.09 (1.39) |
| Sum of trials 1–4 (0–40) | | | | | |
| NEC | 30.25 | −1.34 | 30.77 (64) | 31.74 (62) | 0.97 (3.98) |
| AD I | 17.76 | −1.84 | 18.06 (50) | 13.61 (44) | −4.45 (5.92) |
| AD II | 14.12 | −2.63 | 14.06 (47) | 9.18 (40) | −4.89 (4.27) |
| AD III | 5.79 | −.63 | 9.02 (45) | 5.00 (29) | −4.02 (3.80) |
| Delayed recall (0–10) | | | | | |
| NEC | 7.75 | −0.22 | 8.02 (64) | 8.48 (61) | 0.41 (1.16) |
| AD I | 2.40 | −0.16 | 2.30 (50) | 1.45 (44) | −0.89 (1.74) |
| AD II | 1.54 | −0.63 | 1.36 (47) | 0.60 (42) | −0.76 (1.38) |
| AD III | 0.84 | −0.37 | 0.71 (45) | 0.20 (30) | −0.48 (0.99) |

NEC, normal elderly control; AD-I, MMSE ≥ 21; AD-II, MMSE 16–20; AD-III, MMSE 10–15.

chance levels at baseline, whereas AD group I's performance was better. However, substantial numbers of subjects in all three groups performed at chance levels at baseline (12%, 23%, and 35%, respectively). By month 12 the proportion of individuals scoring at chance increased in each group. Of the NEC, 52% scored perfectly at baseline. By month 12, 36.9% of AD patients remaining in the study scored at chance levels, whereas 48% of NEC scored perfectly. For NEC subjects and AD I patients, there was virtually no change over 1 year. The change for AD group II was small, and only in the AD III group was the mean change even close to the size of the standard deviation ($d = 0.89$).

### Cancellation Test of Visual Attention and Concentration

Results for each of the kinds of cancellation task are presented in Table 3. One-month retest reliabilities were excellent ($r > 0.86$) for the Single Letter, Pair of Letters, and Either of Two Numbers tasks, good ($r = 0.90$) for the Either of Two Single Letters task, and slightly less

**TABLE 2.** *Results of the face recognition memory task*

| Group | Baseline: 1-month follow-up subjects | 1-month change score | Baseline: all subjects (*n*) | 12-month score for all subjects (*n*) | 12-month change (SD) |
|---|---|---|---|---|---|
| Number correct (0–12) | | | | | |
| NEC | 11.03 | −0.44 | 11.13 (64) | 11.08 (62) | −0.05 (1.21) |
| AD I | 8.68 | 0.08 | 8.90 (50) | 8.53 (45) | −0.16 (2.54) |
| AD II | 7.04 | 0.29 | 7.68 (47) | 6.90 (41) | −0.88 (2.48) |
| AD III | 7.33 | −0.83 | 7.37 (43) | 6.16 (25) | −1.92 (2.16) |

Abbreviations as in Table 1.

**TABLE 3.** *Results of the letter cancellation task*

| Group | Baseline: 1-month follow-up subjects | 1-month change score | Baseline: all subjects (*n*) | 12-month score for all subjects (*n*) | 12-month change (SD) |
|---|---|---|---|---|---|
| Single Letter "P" (0–40) | | | | | |
| NEC | 30.34 | 7.19 | 30.31 (64) | 31.82 (62) | 1.27 (3.52) |
| AD I | 22.01 | 7.75 | 23.16 (49) | 19.76 (45) | −3.39 (5.34) |
| AD II | 16.00 | 5.46 | 16.30 (47) | 11.95 (37) | −4.49 (6.31) |
| AD III | 8.81 | 6.61 | 10.45 (44) | 6.67 (24) | −3.21 (4.78) |
| Either of Two Single Letters "L" or "T" (0–40) | | | | | |
| NEC | 23.97 | 4.81 | 23.83 (64) | 25.69 (62) | 1.63 (4.57) |
| AD I | 13.68 | 4.16 | 14.42 (50) | 12.40 (45) | −2.22 (4.09) |
| AD II | 10.04 | 1.79 | 10.11 (47) | 7.49 (37) | −2.62 (4.86) |
| AD III | 4.32 | 2.00 | 5.89 (44) | 3.38 (24) | −2.62 (4.04) |
| Pair of Letters "OE" (0–20) | | | | | |
| NEC | 18.09 | −0.47 | 18.19 (64) | 18.08 (62) | −0.26 (2.34) |
| AD I | 11.72 | −1.40 | 11.78 (50) | 9.60 (45) | −1.93 (3.93) |
| AD II | 8.34 | 0.08 | 9.04 (46) | 6.03 (37) | −2.72 (4.05) |
| AD III | 3.64 | 1.22 | 4.67 (42) | 3.50 (22) | −0.91 (1.95) |
| Set of Three Letters "WKM" (0–13) | | | | | |
| NEC | 11.24 | −2.13 | 11.36 (64) | 11.37 (60) | 0.06 (1.37) |
| AD I | 8.20 | −2.92 | 7.94 (50) | 6.11 (45) | −1.76 (4.25) |
| AD II | 5.25 | −0.88 | 4.98 (47) | 3.95 (37) | −1.08 (2.16) |
| AD III | 2.65 | −0.12 | 3.12 (41) | 2.09 (22) | −0.95 (1.43) |
| Large Letter "LARGE LeTters" (0–40) | | | | | |
| NEC | 33.25 | 3.47 | 34.17 (64) | 38.73 (60) | 4.77 (11.05) |
| AD I | 24.84 | −0.76 | 24.76 (50) | 24.80 (45) | 1.16 (11.76) |
| AD II | 17.75 | −3.04 | 19.57 (47) | 13.78 (37) | −6.81 (9.36) |
| AD III | 8.78 | 1.06 | 11.34 (41) | 8.05 (21) | −3.38 (13.14) |
| Either of Two Numbers "2" or "8" (0–40) | | | | | |
| NEC | 27.31 | −0.59 | 27.22 (64) | 28.61 (62) | 1.26 (4.84) |
| AD I | 18.48 | −2.44 | 18.56 (50) | 15.87 (45) | −2.60 (4.72) |
| AD II | 12.79 | 0.46 | 13.62 (47) | 10.35 (37) | −3.41 (4.90) |
| AD III | 7.20 | 0.06 | 8.95 (40) | 5.64 (22) | −3.23 (4.10) |

Abbreviations as in Table 1.

good for the Set of Three Letters ($r = 0.72$) and Large Letters ($r = 0.74$) tasks. There were no significant correlations with age or education of 12-month change on any of these tasks. There appeared to be sizable learning effects over 1 month for the Single Letter and the Either of Two Single Letters tasks, in that month 1 performance was substantially better than baseline for both tasks. It is possible that these differences resulted from the specific targets used at baseline and 1 month. Further analyses of the two test sequences will be necessary to determine whether these changes are due to repeated exposure to the test or to the specific targets used.

Table 3 indicates that all versions of this task differentiated among the four groups, although the overall difficulty level varied by task. Mean performance of the NEC group was highest for the Pair of Letters task and lowest for the version with either of two single letters as targets. The 12-month change scores indicate that performance was either unchanged or improved somewhat in NEC subjects, whereas performance declined in most AD groups on all versions. On balance, it appears that the change scores were most robust in the cancellation task for Either of Two Numbers, and there is little that the other versions of this task added to the data obtained with the Either of Two Numbers version.

### Maze Completion Test of Executive Function

Table 4 presents data on the number of mazes completed by subjects in each group at baseline and at 12 months. Using the total number of mazes completed by each person as a score, the calculated 1-month retest reliability was very high ($r = 0.95$). There were no confounding effects of age or education on 12-month change scores for this measure. As indicated by the data at the bottom of Table 4, there was no substantial change in the mean number of mazes completed at the 1-month retest reliability session for any subject group (one-sample $t$ test $p > 0.35$ in all cases).

As indicated by the data in Table 4, the more difficult mazes, starting with 4 and higher, were difficult enough that even some of the mildly demented AD patients could not complete them, particularly at 1 year. Hence, time measures would be of little use except for mazes

**TABLE 4.** *Results for the maze completion test of executive function*

| | NEC | AD I | AD II | AD III |
|---|---|---|---|---|
| Percent completed at baseline | | | | |
| Maze 1 | 100 | 100 | 100 | 89 |
| Maze 2 | 100 | 100 | 100 | 87 |
| Maze 3 | 100 | 100 | 96 | 76 |
| Maze 4 | 100 | 96 | 81 | 54 |
| Maze 5 | 100 | 92 | 60 | 35 |
| Maze 6 | 98 | 62 | 32 | 26 |
| Maze 7 | 94 | 38 | 28 | 17 |
| Percent completed at month 12 visit | | | | |
| Maze 1 | 97 | 90 | 81 | 54 |
| Maze 2 | 97 | 90 | 81 | 43 |
| Maze 3 | 97 | 88 | 74 | 33 |
| Maze 4 | 97 | 82 | 55 | 22 |
| Maze 5 | 97 | 66 | 38 | 13 |
| Maze 6 | 97 | 46 | 21 | 11 |
| Maze 7 | 92 | 36 | 17 | 07 |

| Group | Baseline: 1-month follow-up subjects | 1-month change score | Baseline: all subjects (*n*) | 12-month score for all subjects (*n*) | 12-month change (SD) |
|---|---|---|---|---|---|
| Number of mazes completed | | | | | |
| NEC | 6.87 | −0.12 | 6.92 (64) | 6.73 (64) | −0.19 (1.10) |
| AD I | 5.76 | −0.20 | 5.88 (50) | 4.98 (50) | −0.90 (2.19) |
| AD II | 4.63 | 0.08 | 4.96 (47) | 3.68 (47) | −1.28 (2.22) |
| AD III | 3.32 | −0.23 | 3.85 (46) | 1.83 (46) | −2.02 (2.33) |

Abbreviations as in Table 1.

1–3, and then only in normals and patients with relatively mild AD. The overall number of mazes completed decreased over 12 months in AD patients but, given the amount of time required to present all mazes, it would not be practical to use this measure in any brief cognitive assessment.

Because so few patients, except for those in the AD-I group, were able to complete any mazes, most of the time to completion data are not useful. Table 5 presents time to completion data for the first (easiest) three mazes for the NEC and the AD-I patients who were able to complete them both at baseline and at 12 months. As the table indicates, NEC subjects were able to complete all three mazes more quickly than the AD-I patients and improved their performance slightly over 12 months. The AD-I patients, by contrast, were slower on all three mazes after 12 months.

**TABLE 5.** *Time scores for maze completion test*

| Group | Baseline | 12-month | 12-month change (SD) |
|---|---|---|---|
| Maze 1 (seconds) | | | |
| NEC ($n$ = 62) | 2.9 | 2.4 | −0.5 (3.9) |
| AD I ($n$ = 45) | 5.3 | 20.7 | 15.4 (49.3) |
| Maze 2 | | | |
| NEC ($n$ = 62) | 5.5 | 5.2 | −0.3 (2.7) |
| AD I ($n$ = 45) | 13.0 | 33.4 | 20.4 (59.7) |
| Maze 3 | | | |
| NEC ($n$ = 62) | 27.2 | 26.6 | −0.7 (11.5) |
| AD I ($n$ = 44) | 53.2 | 69.4 | 16.2 (59.4) |

Abbreviations as in Table 1.

### Praxis

Table 6 presents the mean scores on this test out of a maximal possible score of 24. Excluding NEC subjects, the retest reliability was modest ($r$ = 0.48). However, there was one subject (an AD group III patient) who dramatically declined between baseline and month 1. Because it is likely his disease had progressed within that month and therefore violated the assumption of stability between test sessions, this individual was excluded from the analysis, resulting in a much more favorable retest reliability estimate of $r$ = 0.72. Scores changed little from baseline to 1 month, indicating that practice effects were relatively small for this task, although performance in all but the most severe patients was near ceiling, which would limit the amount of improvement that could be detected. There were no confounding effects of age or education.

In contrast to the measures previously discussed, even AD group II and group III patients performed praxis commands so well at baseline that 48% and 22% of subjects, respectively, scored perfectly. However, by month 12, only 16% and 0% of AD group II and III patients were still performing at ceiling levels. Praxis was easy enough that floor performance occurred for

**TABLE 6.** *Number of correct responses on the praxis tasks*

| Group | Baseline: 1-month follow-up subjects | 1-month change score | Baseline: all subjects | 12-month score for all subjects | 12-month change (SD) |
|---|---|---|---|---|---|
| NEC | 23.5 | 0.03 | 23.9 | 23.8 | −0.10 (0.46) |
| AD I | 23.9 | 0.68 | 23.2 | 22.4 | −0.70 (2.47) |
| AD II | 22.3 | 0.74 | 22.7 | 20.8 | −1.90 (3.15) |
| AD III | 20.1 | 0.89 | 20.5 | 16.7 | −3.90 (4.78) |

Abbreviations as in Table 1.

only one of the 13 AD group III patients still testable at month 12. The NEC group subjects overwhelmingly (more than 90%) performed at ceiling levels at baseline and at 12 months. Performance declined little over the 12 months of follow-up except for the most severe group of AD patients.

## DISCUSSION

The goal of this study was to determine whether any of the present cognitive assessment procedures are likely to be of use in antidementia clinical trials. Ideally, a new measure would be reliable, sensitive to change, different from other standard measures, such those included in the ADAS, and free of floor and ceiling effects. Only one of the cancellation tasks clearly fitted that description entirely. Although not ideal in one respect or another, some of the other measures, such as the delayed recall and the simpler mazes measures, may be of use in selected circumstances.

Of the cancellation tasks the Either of Two Numbers task (task 6) may have been the most favorable, in that it showed no learning, was reliable, and showed declining performance over time in all three patient groups. It did not show ceiling effects at baseline either in the NEC group or the mildest AD group. Although AD group III scored much worse on this task than the milder patients, they were still largely above floor-level performance. Some of the other versions of the cancellation task might also be of use but would probably add little to what could be learned with the two-number task. A desirable feature of a cancellation task is that it requires little extra time to perform. Cancellation presumably requires attention and sustained concentration; it also requires subjects to maintain a cognitive "set" over time in the face of distraction. These cognitive functions are not explicitly represented on the ADAS.

The mazes were also potentially valuable, but the time necessary to administer all forms of the task make it difficult to incorporate into an instrument for treatment trials. The number of mazes completed was highly reliable, and there were no major confounds. There were ceiling effects in the NEC group but not in the AD patients. Maze performance is believed to be one of the prototypical executive functions, involving foresight, planning, maintenance of set, freedom from distraction, and reasoning. Again, those cognitive functions are not explicitly assessed on the ADAS. Because of the low rate of completion of the mazes in all but the mildest AD group the time measures were of little value. Only the times for the three simplest mazes might be of use, and then only in mild AD patients or, possibly, in those "at risk" for AD because of mild cognitive impairment.

The present experimental battery included a 4 trial list learning procedure coupled with delayed recall. Morris et al. (1989) and Welsh et al. (1994) have previously reported data for 10-word three-trial learning on CDR = 1 patients with 1-year follow-up. Our data offer a more refined view of the mildest patients (MMSE scores $\geq$21), who performed better than CDR = 1 patients. The present results demonstrate the reliability and responsiveness of list learning, especially in more mildly affected patients. Furthermore, the present results show that assessment of delayed recall over 1 year is feasible in mild AD patients. The fourth learning trial in verbal learning did not add much directly in terms of reliability or effect size to the present form of the ADAS, which uses three trials, but may have reduced the number of instances of floor performance. Delayed recall is the most sensitive measure for detecting early AD (Welsh et al., 1991). Longitudinal assessment of learning and memory with the enhanced procedures used here may be most important in mild AD patients and in patients with cognitive impairment without dementia (Morris et al., 1991; Locasio et al., 1995; Petersen et al., 1995), rather than in clinical trials that involve more severely affected patients. The addition of delayed recall and, possibly, the addition of a subset of mazes would help to make the ADAS more sensitive in mild cases, for which the current version is not particularly sensitive to change (Stern et al., 1994).

Praxis was unique among the functions we assessed in its suitability for measurement of the more demented patients. It was of little value in milder patients because of the likelihood of achieving perfect scores initially and not declining over 1 year. Although praxis might play a role in studies restricted to more advanced patients (patients with MMSE scores ≤15), it is not suitable for inclusion in an assessment battery intended for milder patients.

Facial recognition was a disappointment, in that a sizable proportion of AD groups II and III performed near chance levels at baseline. Furthermore, in AD group I there was virtually no change over 1 year. Because trials of some antidementia drugs are designed to detect effects on symptom progression rather than symptomatic improvement, the lack of sensitivity of this procedure suggests that it should not be used in clinical trials.

## RECOMMENDATIONS FOR FUTURE COGNITIVE ASSESSMENT

On the basis of these results, the simplest addition to an instrument such as the ADAS would be the Either of Two Numbers cancellation task. It would add only a few minutes to ADAS administration and it could be scored by scaling the number of responses between 1 and 5. Based on the data from the four groups included in this study, the most useful scoring rule would be: ≥30 correct = 0; 24–30 correct = 1; 18–23 correct = 2; 12–17 correct = 3; 6–11 correct = 4; 0–5 correct = 5).

Inclusion of a cancellation task should not change the psychometric properties of the rest of the ADAS, although it will alter calculations of effect size based on prior work with the ADAS. A recent study that analyzed the predictive effects of different cognitive measures on instrumental ADL function emphasized visuoperceptual performance (in that case Poppelreuter's figures of overlapping objects) as being most useful (Hill et al., 1995). Adding cancellation, which has a visual perceptual component, may further strengthen the ADAS in this cognitive domain.

For clinical trials involving AD patients with MMSE scores >20 or involving cognitively impaired, nondemented individuals, four-trial learning plus delayed recall could be added to the ADAS-cog. For trials of "at risk" or very mild AD patients, one or two of the simplest mazes might be a useful addition.

For clinical trials involving AD patients with MMSE scores ≤15, praxis could be added to the battery. Further praxis items are probably not needed, however, because the Severe Impairment Battery (see Schmitt et al., this issue) includes praxis items.

The present attempts to bring measures of nonverbal memory (face recognition) into clinical trials assessments were not successful. Further work must be done to find procedures that are reliable and responsive.

**Acknowledgment:** This work was supported by NIH grant U01 AG 10483. We thank all participating sites of the ADCS for contributing time and effort to enroll and follow patients for this study.

## REFERENCES

Christensen KJ, Mutchup KS, Norstrom S, et al. A cognitive battery for dementia. Development and measurement characteristics. *Psychol Assessment* 1991;3:168–74.

Ferris SF, Mackell JM, Mohs R. A multicenter evaluation of new treatment efficacy instruments for Alzheimer's disease clinical trials: overview and general results. *Alzheimer Dis Assoc Disord* 1997;11(suppl 2):S1–12 (this issue).

Folstein MF, Folstein SE, McHugh PR. "Mini-mental state": a practical method for grading the cognitive status of patients for the clinician. *J Psychiatr Res* 1975;12:189–98.

Goodglass H, Kaplan E. *The assessment of aphasia and related disorders.* Philadelphia: Lea & Febiger, 1972.

Hill RD, Backman L, Fratiglioni F. Determinants of functional abilities in dementia. *J Am Geriatr Soc* 1995;43:1092–7.

Hughes CP, Berg L, Danziger WL, Coben LA, Martin RL. A new clinical scale for the staging of dementia. *Br J Psychiatry* 1982; 140:566–72.

Lezak M. *Neuropsychological assessment.* 2nd ed. New York: Oxford University Press, 1982.

Locasio JJ, Growdon JH, Corkin S. Cognitive test performance in detecting, staging, and tracking Alzheimer's disease. *Arch Neurol* 1995;52:1087–99.

Morris JC. The Clinical Dementia Rating (CDR): current version and scoring rules. *Neurology* 1993;43:2412–4.

Morris JC, Heyman A, Mohs RC, et al. The Consortium to Establish a Registry for Alzheimer's Disease (CERAD). Part I. Clinical and neuropsychological assessment of Alzheimer's disease. *Neurology* 1989;39:1159–65.

Morris JC, McKeel DW, Storandt M, et al. Very mild Alzheimer's disease: informant-based clinical, psychometric and pathological distinction from normal aging. *Neurology* 1991;41:469–78.

Neisser U. Visual search. *Sci Am* 1964;210:94–102.

Pavio A, Yuille JC, Madigan SA. Concreteness, imagery, and meaningfulness values for 925 nouns. *J Exp Psychol* 1968;76:1–25.

Petersen RC, Smith GE, Ivnik RJ, et al. Apolipoprotein E status as a predictor of the development of Alzheimer's disease in memory-impaired individuals. *JAMA* 1995;273:1274–8.

Reisberg B, Ferris SH, deLeon MJ, Crook T. The Global Deterioration Scale for assessment of primary degenerative dementia. *Am J Psychiatry* 1982;139:1136–9.

Rosen WG, Mohs RC, Davis KL. A new rating scale for Alzheimer's disease. *Am J Psychiatry* 1984;141:1356–64.

Stern RG, Mohs RC, Davidson M, et al. A longitudinal study of Alzheimer's disease: measurement, rate and predictors of cognitive deterioration. *Am J Psychiatry* 1994;151:390–6.

Thorndyke EL, Lorge I. *The teacher's wordbook of 30,000 words.* New York: Teachers College, Bureau of Publications, 1944.

Warrington EK, James M. Disorders of visual perception in patients with localized cerebral lesions. *Neuropsychologia* 1967;5:253–66.

Welsh K, Butters N, Hughes J, et al. Detection of abnormal memory decline in mild cases of Alzheimer's disease using CERAD neuropsychological measure. *Arch Neurol* 1991;48:278–81.

Welsh KA, Butters N, Mohs RC, et al. The Consortium to Establish a Registry for Alzheimer's Disease (CERAD). Part V. A normative study of the neuropsychological battery. *Neurology* 1994;44:609–14.

Yesavage JA, Tinklenberg JR, Hollister LE, Berger PA. Vasodilators in senile dementia: a review of the literature. *Arch Gen Psychiatry* 1979;36:220–3.