Use of a false recognition paradigm in an Alzheimer's disease clinical trial: A pilot study

Andrew E. Budson, MD
Kalina J. Michalska, BS
Dorene M. Rentz, PsyD
Claire C. Joubert, PAC
Kirk R. Daffner, MD
Daniel L. Schacter, PhD
Reisa A. Sperling, MD

Abstract

We report the first use of a false recognition memory test in a clinical trial of patients with Alzheimer's disease (AD). Tests of false recognition allow measurement of two components of memory: the specific details of a prior encounter with a particular item (item-specific recollection) and the general meaning, idea, or gist conveyed by a collection of items (gist memory). We used a false recognition paradigm with categorized pictures to study the effects of an experimental medication in patients with AD. Because medications to treat AD may preferentially improve gist memory or item-specific recollection, use of this type of paradigm may improve sensitivity for detection of drug effects more than standard memory tests.

Key words: Alzheimer's disease, memory, Neotrofin

Andrew E. Budson, MD, Assistant Professor of Neurology, Harvard Medical School, Boston, Massachusetts and Associate Neurologist, Division of Cognitive and Behavioral Neurology, Brigham and Women's Hospital, Boston, Massachusetts.

Kalina J. Michalska, BS, Research Assistant, Division of Cognitive and Behavioral Neurology, Brigham and Women's Hospital, Boston, Massachusetts.

Dorene M. Rentz, PsyD, Instructor in Neurology, Harvard Medical School, Boston, Massachusetts and Senior Research Neuropsychologist, Division of Cognitive and Behavioral Neurology, Brigham and Women's Hospital, Boston, Massachusetts.

Claire C. Joubert, PAC, Study Coordinator, Department of Neurology, Brigham and Women's Hospital, Boston, Massachusetts.

Kirk R. Daffner, MD, Assistant Professor of Neurology, Harvard Medical School, Boston, Massachusetts and Chief, Division of Cognitive and Behavioral Neurology, Brigham and Women's Hospital, Boston, Massachusetts.

Daniel L. Schacter, PhD, Professor and Chair, Department of Psychology, Harvard University, Cambridge, Massachusetts.

Reisa A. Sperling, MD, Instructor in Neurology, Harvard Medical School, Boston, Massachusetts and Director of Clinical Trials for Alzheimer's Disease, Division of Cognitive and Behavioral Neurology, Brigham and Women's Hospital, Boston, Massachusetts.

Introduction

Patients with probable Alzheimer's disease (AD) not only fail to retrieve desired information but also suffer from distortions of memory.[1] These memory distortions may impair the ability of the patient with AD to live independently.[2] For example, patients may believe that they turned off the stove or took their medication when they only thought about performing these activities.

False recognition is a type of memory distortion that occurs when participants falsely report that they have seen non-studied items that are in some way related to studied items. For example, after studying lists of semantically related words (*e.g.*, candy, sour, sugar, bitter, good, taste, etc.) that all converge on a non-presented "theme word" or "related lure" (*e.g.*, sweet), participants frequently intruded the related lure on free-recall tests[3]

and made very high levels of false alarms to these words on recognition tests.[4] Schacter and colleagues[5] have suggested that true and false recognition in this type of paradigm depend on memory for two different kinds of information: specific details of a prior encounter with a particular item (item-specific recollection) and the general meaning, idea, or gist conveyed by a collection of items (gist information; c.f., Reyna and Brainerd,[6] Schacter, Norman, and Koutstaal[7]). As the individual items are presented in this type of paradigm, a gist representation is developed, which may result in an experience of recollection or familiarity when either a studied item or a related lure is presented on a later recognition test. Thus, in this type of paradigm, accurate recognition of previously studied items probably depends on both item-specific and gist information, whereas false recognition of related lure words depends on remembering gist but not item-specific information.[8-10]

Recently, patients with mild to moderate AD have been examined with several false recall and false recognition paradigms using semantically related words,[11] phonologically related words,[12] and other types of stimuli. In one false recognition experiment, using repeated study-test trials of semantically related words, item-specific recollection was found to be the most sensitive measure that distinguished patients with mild AD from healthy older adults.[13] In addition, although patients with AD were unable to improve their poor item-specific recollection, in this experiment they were able to increase their gist memory by repeated study-test trials.

Using a different false recognition paradigm consisting of perceptually related novel objects, both patients with AD and older adults demonstrated little use of item-specific recollection.[14] In this setting, patients with AD demonstrated degraded gist memory compared to healthy older adults. In a third investigation, it was found that studying pictures along with words allowed patients with AD to build up greater levels of gist memory than studying words alone; in contrast, no change in item-specific recollection was observed.[15] Following up on this finding, patients with AD were tested with a false memory paradigm consisting of categories of colored photographs. In this categorized pictures study, patients with AD demonstrated significant levels of both gist memory and item-specific recollection. As expected, gist memory was augmented and item-specific recollection was diminished with increasing numbers of categorized items seen during the study session.[16]

These studies illustrate that using different paradigms, patients with AD may show impairments in gist memory, item-specific recollection, or both; in addition, these experimental manipulations demonstrate that patients with AD can show improvements in either one or both of these memory components.[13-16] Impairment in gist memory is typically seen after damage to medial temporal lobes,[5,10,17-19] whereas impairment in item-specific recollection may be observed after damage to either medial temporal or frontal lobes.[5,20-22] Since the pathology of AD affects both medial temporal and frontal lobes,[23,24] it is not surprising that patients with AD exhibit deficits in both of these components of memory.

Thus, a treatment for AD that improves the function of the medial temporal lobes alone may preferentially increase gist memory. Conversely, a treatment that improves the function of frontal lobes alone may preferentially increase item-specific recollection. Treatments that improve the function of both medial temporal and frontal lobes may increase both item-specific recollection and gist memory. Item-specific recollection may, however, be increased to a greater extent than gist memory if the medial temporal and frontal lobes work together synergistically to allow recollection of the specific details of prior encounters with particular items.

We undertook this pilot study to investigate whether a false recognition test that allows separate analyses of gist memory and item-specific recollection could be a more sensitive measure of the benefit to cognition of novel therapeutic agents to treat Alzheimer's disease than conventional memory tests. In contrast, conventional memory tests using unrelated words typically measure the combination of gist memory and item-specific recollection. To our knowledge, this study is the first time a false recognition test has been used in a clinical trial. Thus, we also wanted to determine the feasibility of this type of testing in a clinical trial of patients with AD. Since, compared to other false recognition paradigms, the one using categorized pictures most enabled patients with AD to develop robust levels of both item-specific recollection and gist memory,[16] we used this false recognition paradigm as a single study center add-on to a multi-center phase 2 clinical trial of Neotrofin™ in patients with AD.

Neotrofin (also known as AIT-082) is a hypoxanthine derivative that has been shown to increase a number of growth factors including BDNF, NGF, bFGF, and NT-3, and to stimulate neurite outgrowth *in vitro*.[25,26] In animal studies, Neotrofin has been shown to improve age-induced working memory deficits in mice, to increase mRNA of growth factors in cortex and hippocampus, to increase hippocampal sprouting, and to improve functioning in rat striatum after NMDA-induced damage.[25,27] Because there is evidence that patients with either frontal lobe or medial temporal lobe damage may show particularly poor item-specific recollection,[5,20-22] we suspected that a drug that improved the function of frontal networks (by improving the function of the striatum[28-30]) and hippocampal function (by increasing growth factors

and neurite sprouting in the hippocampus) might be particularly effective at improving item-specific recollection.

In addition to these experimental measures of gist memory and item-specific recollection, we also present data from the conventional memory test performed during this study, the Alzheimer's Disease Assessment Scale-cognitive subscale (ADAS-cog[31]). This test has been widely used in previous clinical trials of experimental medications in patients with AD.[32-35] Besides the total ADAS-cog score, we also looked at two memory tests sub-scales: *Word Recall* and *Word Recognition*.

We hypothesized that we might detect changes in item-specific recollection if Neotrofin preferentially increased this component of memory. We realized, however, that with the small number of participants in this pilot study we would be unlikely to find significant differences even if Neotrofin increased item-specific recollection substantially due to lack of sufficient power. Thus, the primary goal of this study was to determine the feasibility of false recognition testing in a clinical trial of patients with AD.

Methods

Participants

Sixteen patients with a clinical diagnosis of probable AD (National Institute of Neurological and Communicative Disorders and Stroke-Alzheimer's Disease and Related Disorders Association criteria used[36]) who were participating in a separate double-blind, placebo controlled trial of the drug Neotrofin at Brigham and Women's Hospital, Boston, Massachusetts, USA, were recruited for this study. The patients were in the mild to moderate stage of dementia based upon their Mini Mental Status Examination (MMSE[37]) scores (Table 1). A separate written informed consent was obtained from all participants and their caregivers. The study was approved by the human subjects committee (institutional review board) of Brigham and Women's Hospital, and by NeoTherapeutics, the producer of Neotrofin. Participants were excluded if they were characterized by clinically significant depression, alcohol or drug use, brain damage, or if English was not their primary language. All participants had normal or corrected to normal vision.

Procedure

Participants were randomly assigned to receive either Neotrofin 150 mg or placebo for 90 days. Participants were individually tested at day 0 (baseline) and day 90. The experimental stimuli were divided into two equal sets in order to expose participants to different categories of items during the two sessions. The set of items seen on the first versus the second testing session was counterbalanced. The ADAS-cog and the MMSE were tested as part of the standard Neotrofin protocol.

All participants were tested individually in the clinical trials center at Brigham and Women's Hospital, Boston, Massachusetts, USA. The experimental procedure took between 25 and 40 minutes and involved three phases: a study phase, a brief retention interval, and a test phase.

In the study phase, participants were presented with each item for two seconds, and were asked to rate their liking ("like" or "dislike") for each of them. Although each picture disappeared after two seconds, the liking rating was self-paced. The pictures from different categories were randomly intermixed and the encoding task was incidental—no mention was made of a subsequent memory test. Participants stated their liking rating orally and the experimenter then entered the appropriate response on the keyboard. During the brief five-minute retention interval, participants performed simple puzzles.

In the test phase, participants were given a surprise recognition test and were asked to designate each item as "old" (previously presented during the study phase) or "new" (not previously presented during the experiment). Following the test phase, which was self-paced, participants were debriefed.

Stimuli

The stimuli were identical to those used in Koutstaal *et al.*,[17] and consisted of colored photographs of single objects (or, in a few cases, coherent grouping of objects), without background, taken from various illustrated books for children and adults. All pictures were initially mounted on white index cards and then scanned and converted to digital format using VistaScan and a UMAX Vista-S6E scanner (UMAX Technologies, Inc., 3561 Gateway Blvd., Fremont, CA 94538 USA). At both study and test, the pictures were displayed in the center of a color computer monitor using an Apple Macintosh Powerbook 5300c computer and PsyScope software.[38]

The pictures portrayed objects from 25 different categories (*e.g.*, birds, motorcycles, toys), each with 21 different exemplars. There were also 30 pictures of unrelated objects. The categories were randomly assigned to six sets of four categories each (one of the 25 categories, chosen randomly, was not used), and each set was rotated through the experimental conditions such that each set equally often served as a study category comprised of three or 18 related items or as a novel, non-studied item. When a given category served as an 18-exemplar category, 18 of the 21 items were presented at study; the remaining three

Table 1. Demographics, mini mental status examination, and Alzheimer's disease assessment scale-cognitive subscale

		Age (years)	Education (years)	MMSE	ADAS-cog* Total	ADAS-cog* Word recall	ADAS-cog* Word recognition
Day 0							
Neotrofin	*M (SD)*	70.6 (10.0)	17.4 (3.8)	21.9 (2.54)	22.6 (5.93)	7.3 (1.10)	5.0 (2.06)
	F (1,15), p	< 1, *ns*	< 1, *ns*	< 1, *ns*	2.68, *ns*	< 1, *ns*	2.81, *ns*
Placebo	*M (SD)*	72.8 (7.8)	15.6 (4.1)	21.3 (3.04)	27.1 (5.04)	7.2 (1.45)	7.0 (2.57)
Day 90							
Neotrofin	*M (SD)*			22.1 (3.44)	23.7 (6.68)	7.1 (1.38)	5.6 (2.93)
	F (1,15), p			< 1, *ns*	1.24, *ns*	< 1, *ns*	1.91, *ns*
Placebo	*M (SD)*			20.1 (4.78)	29.0 (11.13)	7.3 (1.63)	7.3 (1.97)
Change							
Neotrofin	*M*			0.3	1.1	-0.2	0.6
	F (1,15), p			< 1, *ns*	< 1, *ns*	< 1, *ns*	< 1, *ns*
Placebo	*M*			-1.2	1.9	0.1	0.3

Note: *ns* = nonsignificant, $p > .10$; Change = Day 90 minus Day 0. *Higher scores on the ADAS-cog indicate worse performance.

items were reserved to be presented during the recognition test as new but related lure items. Likewise, when a given category served as a three-exemplar category, only a fraction of the total pool of items was presented at study. In these latter cases, the particular items excluded were determined randomly, with the same items always excluded whenever that three-exemplar category was presented.

To avoid confounding the number of items per category that were presented at study with the number of items presented at test, three studied and three lure items were presented per category for the 18- and three-exemplar categories. This was accomplished by selecting a subset of items from each category that always served as the critical study and test items. For each category, six items were randomly selected to serve as the critical target and lure items. These items were then assigned to two subsets (A & B) and were rotated through the study and test lists such that each subset served equally often as targets and lures for the studied categories, or as novel items for the non-studied categories. The novel categories were also tested three times. Full counterbalancing required 24 participants per group.

For the unrelated items, 24 of the total 30 items were chosen randomly to be used in the experiment. These were divided into two sets (X & Y) of 12 unrelated items each. Half of the participants were shown set X during the study session; the other half were shown set Y. Both sets were presented during the test session, scored appropriately as either studied or non-studied unrelated items.

Each study list was comprised of a total of 102 items: items from four 18-exemplar categories (4 x 18); four three-exemplar categories (4 x 3); 12 unrelated items (12 x 1); three primacy and three recency buffers. Each test list was comprised of a total of 84 items: three studied and three lure items from each of the 18- and three-exemplar categories (four categories x six items/ category x two category sizes); three novel items each from four non-studied novel categories (3 x 4); and twelve studied and twelve non-studied unrelated items (12 x 2). Study and test lists were presented in pseudo random order.

Statistical analysis

The experimental design included a between-subjects factor of group (Neotrofin vs. placebo) and a within-subjects factor of category size. For studied items, category size had two levels: three and 18 category exemplars presented. Non-studied items had three levels of category

size—the aforementioned two levels plus novel category items for which no related items were present at study. Studied and non-studied unrelated items were also used to provide a measure of participants' performance on a more standard memory test.

The novel items provided an estimate of baseline levels of false alarms. True recognition in this paradigm is comparable to conventional memory tests, since both gist memory and item-specific recollection likely contribute to true recognition. Because false recognition in this paradigm depends on remembering gist but not item-specific information, false recognition provided a measure of gist-based memory. Lastly, since true recognition can be thought of as a combination of gist memory plus item-specific recollection, and false recognition is likely a measure of gist memory minus any item-specific recollection that is available to counteract the effect of gist, subtracting false recognition from true recognition should provide a measure of the item-specific recollection used by the participants.

One-way analyses of variance (ANOVAs) were used to compare the Neotrofin and placebo groups in demographic variables, MMSE, ADAS-cog, novel and unrelated items at days 0, 90, and the change between days 0 and 90. Repeated measures ANOVAs were used to compare the Neotrofin and placebo groups in true recognition, false recognition, and item-specific recollection.

Results

Table 1 presents the results of the demographic data, the MMSE, and the ADAS-cog as a function of test session (day 0 or 90) and group (Neotrofin vs. placebo), as well as the between-group significance. Note that higher scores on the ADAS-cog indicate a decrement in performance. Also shown in Table 1 is the change in test scores from day 0 to day 90. Table 2 presents the proportion of "old" responses to studied items (true recognition) and non-studied items (false recognition) as a function of category size (18-and three-exemplar), test session (day 0 or 90), and group (Neotrofin vs. placebo). Also shown in Table 2 is the proportion of "old" responses to novel items, and to studied and non-studied unrelated items.

Analyses are presented for the demographic data, MMSE and ADAS-cog, unrelated items, novel items (baseline false alarms), true recognition (equivalent to standard memory tests), false recognition (a measure of gist-based memory), and item-specific recollection. Data from day 0 is presented first, followed by day 90, and then finally the change from day 0 to 90. To correct for response bias (participants overall tendency to respond "old" vs. "new") on the experimental data, all analyses were performed on the corrected data by subtracting the proportion of baseline false alarms ("old" responses to novel items) from the uncorrected true and false recognition responses in Table 2. Similarly, unrelated items were corrected by subtracting the proportion of "old" responses to non-studied unrelated items from the "old" responses to studied items.

Demographics

The Neotrofin and placebo groups were well matched with respect to gender (Neotrofin: five women, two men; placebo: seven women, two men), age, education, and day 0 MMSE scores (Table 1).

MMSE and ADAS-cog

As can be seen in Table 1, there were no significant differences in the ADAS-cog Total, sub-tests of the ADAS-cog (Word Recall and Word Recognition), or MMSE between the Neotrofin and placebo groups at days 0 and 90. There were also no significant between-group differences in the change between days 0 and 90 (Table 1).

Unrelated items: Standard memory test

There were differences at day 0 between the groups in their responses to unrelated items, with the placebo group showing a much higher corrected hit rate to unrelated items compared with those in the Neotrofin group ($F(1,14) = 7.41$, $MSE = .066$, $p = .017$) (Table 2). No difference between the groups was seen at day 90 ($F(1,14) < 1$). Reflecting the day 0 difference, however, there was a significant difference between the groups for the day 0 to day 90 change in unrelated items ($F(1,14) = 8.97$, $MSE = .072$, $p = .010$), with the Neotrofin group showing an increase in their level of unrelated items and the placebo group showing a decrease.

Novel items: Baseline false alarms

There were no significant differences between novel false alarms for those in the placebo group compared to those in the Neotrofin group for the day 0 data (Table 2). A one-way ANOVA showed no effect of Group (Neotrofin vs. placebo) ($F(1,14) < 1$). Similarly, there was no effect of Group for either the day 90 data or the change between days 0 and 90 ($Fs(1,14) < 1$).

True recognition: Gist memory plus item-specific recollection

For the day 0 data, a 2 (Group: Neotrofin vs. placebo)

Table 2. True and false recognition responses by category size								
		True recognition			False recognition			
Category size		18	3	unrelated	18	3	novel	unrelated
Day 0								
Neotrofin	*M*	.74	.57	.62	.67	.37	.31	.50
	SD	.18	.30	.17	.21	.26	.18	.29
	F (1,15), p	< 1, *ns*	< 1, *ns*	< 1, *ns*	1.66, *ns*	< 1, *ns*	< 1, *ns*	1.57, .001
Placebo	*M*	.69	.58	.56	.54	.44	.22	.08
	SD	.21	.16	.26	.19	.18	.20	.11
Day 90								
Neotrofin	*M*	.58	.65	.55	.50	.40	.29	.23
	SD	.19	.17	.19	.23	.20	.23	.28
	F (1,15), p	1.37, *ns*	< 1, *ns*	< 1, *ns*	< 1, *ns*	< 1, *ns*	< 1, *ns*	< 1, *ns*
Placebo	*M*	.70	.56	.44	.52	.48	.26	.17
	SD	.22	.23	.26	.23	.25	.13	.17

Note: *ns* = nonsignificant, $p > .10$.

x 2 (Category Size: 18- or three-exemplar) ANOVA showed an effect of Category Size ($F(1,14) = 5.10$, $MSE = .030$, $p = .040$), but no effect of Group ($F(1,14) < 1$) and no Group x Category Size interaction ($F(1,14) < 1$). The effect of category size indicates that, overall, participants showed greater levels of true recognition when more category exemplars were shown during study, consistent with previous work[14,16,17,19] (Table 2).

Analysis of the day 90 data showed no effect of Category Size ($F(1,14) < 1$), no effect of Group ($F(1,14) < 1$), and no Group x Category Size interaction ($F(1,14) = 2.42$, $MSE = .036$, $p = .142$). The analysis of the change in true recognition between days 0 and 90 also yielded no significant effects or interactions (Group: $F(1,14) < 1$; Category Size: $F(1,14) = 1.92$, $MSE = .045$, $p = .188$; Group x Category Size: $F(1,14) = 3.07$, $MSE = .045$, $p = .102$).

False recognition: Gist memory

In the analysis of the data at day 0, an ANOVA showed an effect of Category Size ($F(1,14) = 13.02$, $MSE = .024$, $p = .003$), no effect of Group ($F(1,14) = 1.00$), and no Group x Category Size interaction ($F(1,14) = 3.13$, $MSE = .024$, $p = .099$). The effect of Category Size indicates that participants showed greater levels of false recognition when more category exemplars were shown during study, consistent with previous work[14,16,17,19] (Table 2).

Analysis of the day 90 data demonstrated no effect of Category Size, no effect of Group and no Group x Category interaction (all $Fs(1,14) < 1$). Likewise, the analysis of the change in false recognition between days 0 and 90 showed no effect of Group ($F(1,14) < 1$), no effect of Category Size ($F(1,14) = 2.61$, $MSE = .054$, $p = .129$), and no Group x Category Size interaction ($F(1,14) < 1$).

Item-specific recollection: True recognition minus false recognition

Analysis of the data from day 0, day 90, and the change in item-specific recollection between days 0 and 90, showed no effect of Category Size, no effect of Group, and no interactions (all $Fs(1,14) < 1$ except the day 90 interaction: $F(1,14) = 3.02$, $MSE = .047$, $p = .104$).

Discussion

Previous studies of medications to treat patients with AD have used memory tests that measure the combination of gist memory and item-specific recollection. Tests of false recognition, which allow the separate analysis of gist memory and item-specific recollection, may be more sensitive to effects of treatment in AD with novel therapeutic agents than conventional memory tests. To our knowledge, the present study represents the first time that the effects of a novel medication to treat patients with AD have been evaluated with a false recognition paradigm.

The Neotrofin and placebo groups were well matched with respect to gender, age, education, and baseline scores of the standardized cognitive tests (MMSE and ADAS-cog). No differences were observed between the groups and these standard tests at day 90 or in the change between days 0 and 90 (Table 1).

Differences were observed in our day 0 data for the unrelated items, with the placebo group showing a greater corrected hit rate compared to the Neotrofin group (Table 2). Although there was no difference in the day 90 data, the difference in the day 0 data led to a significant difference between the groups for the change between days 0 and 90 in the corrected hit rate for the unrelated items, showing an increase for the Neotrofin group and a decrease for the placebo group. Why the groups should have differed on this variable on day 0 is unclear. One possibility is that because we had a small number of subjects, individual differences in day 0 values may have produced a large effect. Another possibility concerns the fact that the counterbalancing of the stimuli in this study was not perfect because a full counterbalancing requires 24 participants per group. Thus, the Group differences observed may have been due to the particular items shown to the participants at study and test. In either case, repeating the study with larger numbers of participants—and with complete counterbalancing—should eliminate this problem. Further work will be needed to establish the test-test reliability and the face validity of the paradigm in assessing the cognitive status in patients with AD.

Participants showed greater levels of true and false recognition on day 0 with the 18-exemplar categories compared to the three-exemplar ones, consistent with previous work.[14,16,17,19] No differences were observed between the groups for true or false recognition for day 0, day 90, or the change between days 0 and 90. Since item-specific recollection was calculated from the true and false recognition data, it was not surprising that no differences between groups were observed in the analysis of item-specific recollection.

We have shown that it is feasible to use a false memory paradigm as an outcome measure in a clinical trial of patients with AD. We had hypothesized that we might detect changes in item-specific recollection if Neotrofin preferentially increased this component of memory. Instead we found no differences between the Neotrofin and placebo groups for true recognition, false recognition, or item-specific recollection. This negative finding is perhaps not surprising considering that most studies that have found efficacy in AD have used several hundred patients[32-35] while we only used 16.

In conclusion, the present study is, to our knowledge, the first clinical trial in patients with AD to use a false recognition paradigm to look for drug effects on two separable components of memory: gist memory and item-specific recollection. We have demonstrated that using this type of memory paradigm is feasible within the setting of a clinical trial. Because cognitive enhancing medications used to treat AD may improve gist memory or item-specific recollection preferentially, use of this type of paradigm may allow more sensitive detection of drug effects than standard memory tests.

Authors' notes

Correspondence concerning this article should be addressed to Andrew E. Budson, MD, Division of Cognitive and Behavioral Neurology, Brigham and Women's Hospital, 75 Francis Street, Boston, MA 02115 or via e-mail at abudson@partners.org. *Although the majority of the investigators (AEB, DMR, CCJ, KRD, RAS) of this study were paid by NeoTherapeutics for their work in the Neotrofin multicenter trial, no additional funding was provided for the false memory testing.*

Acknowledgments

This research was supported by the National Institute on Aging grant AG08441, the National Institute of Mental Health grants F32 MH11767 and K23 MH01870, and by the National Institute of Neurological Diseases and Stroke grants NS26980 and K23 NS02189. We thank NeoTherapeutics for allowing us to perform this study. We thank Wilma Koutstaal, Alison Sullivan, Steve Prince, Danielle Unger, and Leonard Scinto for their invaluable help and support.

References

1. Förstl H, Besthorn C, Burns A, *et al.*: Delusional misidentification in Alzheimer's disease: A summary of clinical and biological aspects. *Psychopathology*. 1994; 27: 194-199.
2. Borson S, Raskind MA: Clinical features and pharmacologic treatment of behavioral symptoms of Alzheimer's disease. *Neurology*. 1997; 48: S17-24.
3. Deese J: On the prediction of occurrence of particular verbal intrusions in immediate recall. *J Exp Psychol*. 1959; 58: 17-22.
4. Roediger HL, McDermott KB: Creating false memories: Remembering words not presented in lists. *J Exp Psychol: Learn Mem Cognit*. 1995; 21: 803-814.
5. Schacter DL, Verfaellie M, Anes MD, Racine C: When true recognition suppresses false recognition: Evidence from amnesic patients. *J Cognit Neurosci.* 1998; 10: 668-679.

6. Reyna VF, Brainerd CJ: Fuzzy-trace theory: An interim synthesis. *Learn Individual Differences*. 1995; 7: 1-75.
7. Schacter DL, Norman KA, Koutstaal W: The cognitive neuroscience of constructive memory. *Annu Rev Psychol*. 1998; 49: 289-318.
8. Brainerd CJ, Reyna VF: Fuzzy-trace theory and children's false memories. *J Exp Child Psychol*. 1998; 71: 81-129.
9. Payne DG, Elie CJ, Blackwell JM, Neuschatz JS: Memory illusions: Recalling, recognizing, and recollecting events that never occurred. *J Mem Lang*. 1996; 35: 261-285.
10. Schacter DL, Verfaellie M, Pradere D: The neuropsychology of memory illusions: False recall and recognition in amnesic patients. *J Mem Lang*. 1996; 35: 319-334.
11. Balota DA, Cortese MJ, Duchek JM, *et al.*: Veridical and false memories in healthy older adults and in dementia of the Alzheimer's type. *Cognit Neuropsychol*. 1999; 16: 361-384.
12. Watson JM, Balota DA, Sergent-Marshall SD. Semantic, phonological, and hybrid veridical and false memories in healthy adults and in dementia of the Alzheimer's type. *Neuropsychology*. 2001; 15: 254-267.
13. Budson AE, Daffner KR, Desikan R, Schacter DL: When false recognition is unopposed by true recognition: Gist-based memory distortion in Alzheimer's disease. *Neuropsychology*. 2000; 14: 277-287.
14. Budson AE, Desikan R, Daffner KR, Schacter DL: Perceptual false recognition in Alzheimer's disease. *Neuropsychology*. 2001; 15: 230-243.
15.Budson AE, Sitarski J, Daffner KR, Schacter DL: False recognition of pictures versus words in Alzheimer's disease: The distinctiveness heuristic. *Neuropsychology*. 2002; 16: 163-173.
16. Budson AE, Sullivan AL, Michalska KJ, *et al.*: Phonological and pictorial false recognition in Alzheimer's disease. Unpublished.
17. Koutstaal W, Verfaellie M, Schacter DL: Recognizing identical vs. similar categorically related common objects: Further evidence for degraded gist-representations in amnesia. *Neuropsychology*. 2001; 15: 268-289.
18.Schacter DL, Verfaellie M, Anes MD: Illusory memories in amnesic patients: Conceptual and perceptual false recognition. *Neuropsychology*. 1997; 11: 331-342.
19.Koutstaal W, Schacter DL, Verfaellie M, *et al.*: Perceptually-based false recognition of novel objects in amnesia: Effects of category size and similarity to category prototypes. *Cognit Neuropsychol*. 1999; 16: 317-341.
20. Parkin AJ, Bindschaedler C, Harsent L, Metzler C: Pathological false alarm rates following damage to the left frontal cortex. *Brain Cogn*. 1996; 32: 14-27.
21. Rapcsak SZ, Reminger SL, Glisky EL, *et al.*: Neuropsychological mechanisms of false facial recognition following frontal lobe damage. *Cognit Neuropsychol*. 1999; 16: 267-292.
22. Schacter DL, Curran T, Galluccio L, *et al.*: False recognition and the right frontal lobe: A case study. *Neuropsychologia.* 1996; 34: 793-808.
23. Lidstrom AM, Bogdanovic N, Hesse C, *et al.*: Clusterin (apolipoprotein J) protein levels are increased in hippocampus and in frontal cortex in Alzheimer's disease. *Exp Neurol*. 1998; 154: 511-521.
24. Price JL, Morris JC. Tangles and plaques in nondemented aging and "preclinical" Alzheimer's disease. *Ann Neurol*. 1999; 45: 358-368.
25. Glasky AJ, Glasky MS, Ritzmann RF, Rathbone MP: AIT-082, a novel purine derivative with neuroregenerative properties. *Expert Opinion Invest Drugs*. 1997; 6: 1413-1417.
26. Middlemiss PJ, Glasky AJ, Rathbone MP, *et al.*: AIT-082, a unique purine derivative, enhances nerve growth factor mediated neurite outgrowth from PC12 cells. *Neurosci Lett*. 1995; 199: 131-134.
27. Glasky AJ, Melchior CL, Pirzadeh B, *et al.*: Effect of AIT-082, a purine analog, on working memory in normal and aged mice. *Pharmacol Biochem Behav*. 1994; 47: 325-329.
28. Degos JD, da Fonseca N, Gray F, Cesaro P: Severe frontal syndrome associated with infarcts of the left anterior cingulategyrus and the head of the right caudate nucleus. A clinico-pathological case. *Brain*. 1993; 116: 1541-1548.
29. McDowell S, Whyte J, D'Esposito M: Differential effect of a dopaminergic agonist on prefrontal function intraumatic brain injury patients. *Brain.* 1998; 121: 1155-1164.
30. Sawaguchi T, Goldman-Rakic PS: D1 dopamine receptors in prefrontal cortex: Involvement in working memory. *Science*. 1991; 251: 947-950.
31. Rosen WG, Mohs RC, Davis KL: A new rating scale for Alzheimer's disease. *Am J Psychiatry*. 1984; 141: 1356-1364.
32. Forette F, Anand R, Gharabawi G: A phase II study in patients with Alzheimer's disease to assess the preliminary efficacy and maximum tolerated dose of rivastigmine (Exelon). *Eur J Neurol*. 1999; 6: 423-429.
33. Raskind MA, Cyrus PA, Ruzicka BB, Gulanski BI: The effects of metrifonate on the cognitive, behavioral, and functional performance of Alzheimer's disease patients. *J Clin Psychiatry*. 1999; 60: 318-325.
34. Rogers SL, Friedhoff LT, the Donepezil Study Group: The efficacy and safety of Donepezil in patients with Alzheimer's disease: Results of a US multicentre, randomized, double-blind, placebo-controlled trial. *Dementia*. 1996; 7: 293-303.
35. Tariot PN, Solomon PR, Morris JC, *et al.*: A 5-month, randomized, placebo-controlled trial of galantamine in AD. *Neurology*. 2000; 54: 2269-2276.
36. McKhann G, Drachman D, Folstein M, *et al.*: 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*. 1984; 34: 939-944.
37. Folstein MF, Folstein SE, McHugh PR: A practical method for grading the cognitive state of patients for the clinician. *J Psychiatr Res*. 1975; 12: 189-198.
38. Cohen JD, MacWhinney B, Flatt M, Provost J: PsyScope: A new graphic interactive environment for designing psychology experiments. *Behav Res Methods Instrum Comput*. 1993; 25: 257-271.