35. Cupples LA, D'Agostino RB, Anderson K, et al. Comparison of baseline and repeated measure covariate techniques in the Framingham Heart Study. Stat Med 1988;7:205–218.
36. D'Agostino RB, Lee ML, Belanger AJ, et al. Relation of pooled logistic regression to time dependent Cox regression analysis: the Framingham Heart Study. Stat Med 1990;9:1501–1515.
37. Alter M, Yamoor M, Harshe M. Multiple sclerosis and nutrition. Arch Neurol 1974;31:267–272.
38. Hernán MA, Olek MJ, Ascherio A. Cigarette smoking and incidence of multiple sclerosis in women. Am J Epidemiol 2001;154 (in press).
39. Hu FB, Stampfer MJ, Rimm E, et al. Dietary fat and coronary heart disease: a comparison of approaches for adjusting for total energy intake and modeling repeated dietary measurements. Am J Epidemiol 1999;149:531–540.
40. National Research Council, Subcommittee on the Tenth Edition of the RDAs, Food and Nutrition Board, Commission on Life Sciences. Vitamin C. Recommended dietary allowances. 10th ed. Washington, DC: National Academy Press, 1989:115–124.
41. National Research Council, Subcommittee on the Tenth Edition of the RDAs, Food and Nutrition Board, Commission on Life Sciences. Vitamin E. Recommended dietary allowances. 10th ed. Washington, DC: National Academy Press, 1989:99–107.

Testosterone supplementation improves spatial and verbal memory in healthy older men

M.M. Cherrier, PhD; S. Asthana, MD; S. Plymate, MD; L. Baker, PhD; A.M. Matsumoto, MD; E. Peskind, MD; M.A. Raskind, MD; K. Brodkin, MD; W. Bremner, MD; A. Petrova, BS; S. LaTendresse, BA; and S. Craft, PhD

Article abstract—*Objective:* To determine the relationship between exogenous testosterone administration and cognitive abilities in a population of healthy older men. *Background:* Serum levels of total and bioavailable testosterone gradually decrease with age in men and are associated with reductions in muscle mass, osteoporosis, decreased sexual activity, and changes in cognition. *Methods:* Twenty-five healthy, community-dwelling volunteers, aged 50 to 80 years, completed a randomized, double-blind, placebo-controlled study. Participants received weekly intramuscular injections of either 100 mg testosterone enanthate or placebo (saline) for 6 weeks. Cognitive evaluations were conducted at baseline, week 3, and week 6 of treatment by use of a battery of neuropsychologic tests. *Results:* Circulating total testosterone was raised an average of 130% from baseline at week 3 and 116% at week 6 in the treatment group. Because of aromatization of testosterone, estradiol increased an average of 77% at week 3 and 73% at week 6 in the treatment group. Significant improvements in cognition were observed for spatial memory (recall of a walking route), spatial ability (block construction), and verbal memory (recall of a short story) in older men treated with testosterone compared with baseline and the placebo group, although improvements were not evident for all measures. *Conclusions:* The results suggest that short-term testosterone administration enhances cognitive function in healthy older men. However, it remains unclear whether these improvements in cognition are attributable to increased testosterone or estradiol levels, or both. The potential role of testosterone vs its metabolites on cognition requires further research.

NEUROLOGY 2001;57:80–88

Serum levels of total testosterone, bioavailable testosterone (testosterone that is not bound to sex hormone–binding globulin), and free testosterone decrease with age in men.[1,2] This gradual decrease is associated with decreased muscle mass and strength, osteoporosis, reduced sexual activity, and changes in cognition.[3] Testosterone replacement therapy in normal older men has shown benefits on body composition, bone mass, muscle strength, and sexual functioning.[4] In addition to peripheral physiologic effects, age-related declines in testosterone levels appear to affect spatial memory. Aging mice show a progressive impairment of spatial learning and memory related to decreases in plasma testosterone, which can be reversed with testosterone administration.[5] In healthy older men, declines in endogenous testosterone levels have been found to correlate significantly with declines in both visual and verbal

From the Departments of Psychiatry and Behavioral Sciences (Drs. Cherrier, Baker, Peskind, Raskind, A. Petrova, and S. LaTendresse) and Medicine (Drs. Bremner, Asthana, Plymate, Matsumoto, and Brodkin), University of Washington Medical School, Seattle; Geriatric Research, Education, and Clinical Center (Drs. Cherrier, Asthana, Plymate, Baker, Matsumoto, Peskind, Brodkin, and Craft) and Mental Illness Research, Education, and Clinical Center (Drs. Peskind and Raskind), Veterans Administration Puget Sound Health Care System, Seattle, WA.

Supported in part by NIA award #K01 AG00858.

Received August 3, 2000. Accepted in final form April 3, 2001.

Address correspondence and reprint requests to Dr. Monique M. Cherrier, Department of Psychiatry, Box 356560, University of Washington Medical School, 1959 NE Pacific, Seattle, WA 98195; e-mail: cherrier@u.washington.edu

memory and verbal fluency (r = 0.53, 0.52, and 0.45).[6] However, studies examining the effects on memory of exogenous testosterone administration in older men have produced mixed results. A double-blind, placebo-controlled study of older, hypogonadal men given biweekly injections of 200 mg testosterone cypionate found improvements in grip strength without significant changes on memory measures.[7] Similarly, a study using a single, 250-mg testosterone enanthate injection found no beneficial effects on cognition assessed 1 week after the injection in a group of healthy older men.[8] However, it is not clear from these studies whether testosterone levels were robustly raised from baseline. In contrast, a study of weekly 150-mg testosterone enanthate injections for 4 weeks reversed spatial working memory deficits in healthy older men.[9] Findings from these previous studies are mixed but suggest that testosterone administration in older men may have beneficial effects on cognition and, in particular, effects on spatial memory.

In this study, we explored the relationship between exogenous testosterone administration and spatial and verbal memory in a healthy older men population by using a moderate replacement dose of testosterone. We hypothesized that testosterone administration would have a selective but beneficial effect on memory improving both verbal and spatial memory but would not necessarily affect language or attention skills. A randomized, placebo-controlled, double-blind design was used, with 6 weeks of treatment followed by 6 weeks of washout. The study design expanded on previously reported studies in several areas. First, cognitive assessments were performed within 24 to 48 hours after injection to capture peak levels of testosterone. The 100-mg weekly testosterone enanthate dose reached a peak level within 24 to 48 hours after injection and gradually decreased to baseline or below baseline 7 days after injection.[10] Several previous studies have assessed cognitive response to hormone manipulation during trough hormone levels, when effects may be minimal or absent. Second, because of possible rapid effects of hormones, cognitive assessments were included early in the treatment period at week 3 of treatment in addition to week 6. Third, to assess possible selective effects of androgens on spatial memory, a psychometric battery was constructed to measure a wide range of cognitive domains. Measures of spatial memory, verbal memory, attention, and language were included. This design allowed for the assessment of change on cognitive domains expected to change (e.g., spatial memory) along with cognitive domains that were predicted to be unaffected (e.g., attention) as a control. No study has attempted to measure a wide range of cognitive domains at various time points of treatment while capturing peak circulating hormone levels. Finally, some previous studies have used the same test to assess change from baseline, thus increasing the potential for learning or practice effects to occur. Only tests suitable for assessment over multiple times (i.e., tests with multiple, comparable versions) were included in this study.

Methods. *Participants.* Participants were healthy older men between the ages of 50 and 80 years recruited from the community through flyers. The study protocol was approved by the University of Washington Institutional Review Board, and approved informed consent procedures were followed. Seventy-seven potential participants who responded to posted flyers were given a brief phone screen. Participants who were eligible according to the phone screen were asked to come in for a more thorough screening visit, including a physical examination and psychiatric and laboratory evaluation to exclude any relevant physical or medical illness. This included tests of liver function, hypertension, and prostate disease (prostate-specific antigen [PSA] level), and digital rectal examination (DRE). Participants with abnormal findings were excluded from the study. Participants also underwent a cognitive screening examination consisting of the Mini-Mental State Examination (MMSE) and the Mattis Dementia Rating Scale (DRS) to ensure that participants were within the normal range for their age.[11,12] Participants with scores at or below the recommended cutoff score (MMSE 26 or lower, DRS 130 or lower) were excluded from the study. Participants with a history of significant alcohol abuse, psychiatric illness, head injury with loss of consciousness, or who were taking medications with any CNS effects or medications, such as cimetidine, that block the androgen receptor were excluded. Participants with previous or current prostate cancer, elevated PSA levels, history of myocardial infarction, abnormal renal or hepatic disease, sleep apnea, previous testosterone treatment, or other gonadal endocrine disorders were also excluded (figure 1).

Procedures. Study design. Participants who were not excluded from screening criteria were randomly assigned to the treatment group or placebo group. An examination of previous hormone manipulation studies in a power analysis with medium effect size (0.35) and power (0.80) indicated that a sample size of approximately 13 subjects per group (treatment and placebo) would be sufficient to detect changes in a mixed-model analysis of variance (ANOVA) with two groups and two "on treatment" assessments. Assignment for each consecutively enrolled participant was made by using a predetermined assignment sheet that was created with a random number generator. Psychometrists who performed the cognitive testing were blinded to treatment conditions. Nurses who administered the injections were not blinded and were responsible for preparing syringes in advance to ensure blinded condition for subjects. Investigators were blinded to condition.

Twenty-five healthy older men who met screening criteria reported to the Special Studies Unit of the Veterans Administration Puget Sound Health Care System (VAPSHCS) for weekly intramuscular injections of 100 mg testosterone enanthate (Delatestryl, manufactured for BTG Pharmaceuticals Corp. by Bristol-Myers Squibb, Princeton, NJ) or placebo (saline). At these weekly visits, blood samples were taken to measure serum testosterone and estradiol levels by IFMA and radioimmunoassay (RIA) (see below). Cognitive testing was conducted at baseline and repeated at weeks 3 and 6 of treatment. Testing ses-

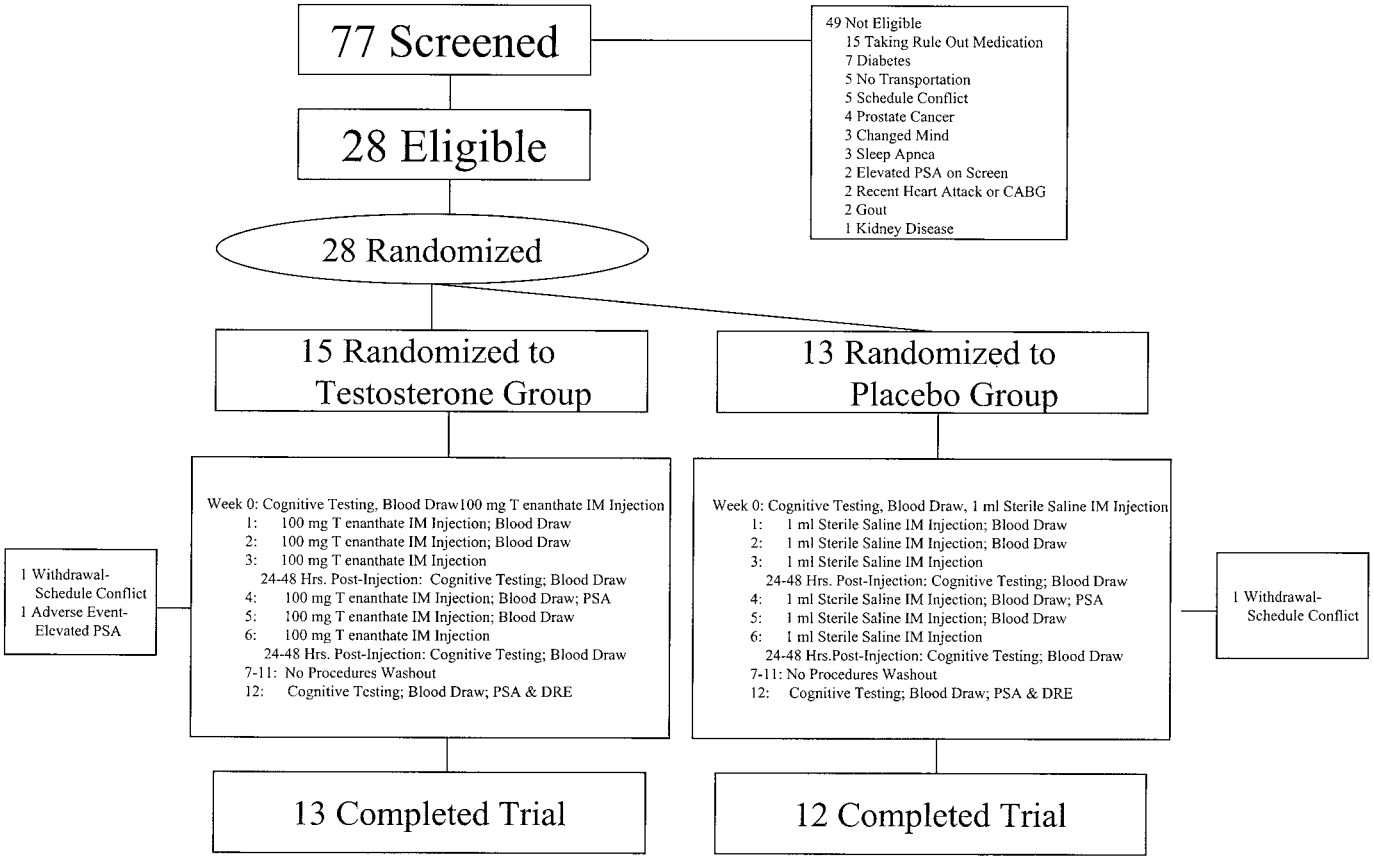

Figure 1. Participant flow and study design sheet.

sions occurred within 24 to 48 hours after testosterone or placebo injection to capture peak testosterone levels for the treated group. Therefore, cognitive performance results reflect the effects of peak testosterone levels. Thirteen participants were randomized into the treatment group, and 12 participants were randomized into the placebo or control group. See figure 1 for study procedures. The treatment and placebo groups did not differ with regard to demographic variables (age and education). Endogenous testosterone levels measured at baseline, before the start of the study, were in the low normal range for healthy older men (15 to 20 nmol/L) and were not significantly different between groups. Normal range for young men is 22 ± 4 nmol/L. PSA levels were measured at screen, at week 4 of the study, and again at washout. The mean PSA level of 1.54 ng/mL for all participants at the start of the study was within normal limits (0 to 4 ng/mL). DRE was also conducted at screen and at washout visit, with no abnormal findings (table 1).

Table 1 *Demographic information*

Demographics	T 100 mg/wk (n = 13)	Placebo (n = 12)
Age, y	65 (9)	70 (7)
Education, y	15 (3)	16 (3)
Dementia Rating Scale, total score*	140 (3)	140 (4)
Baseline total testosterone, nmol/L	20 (6)	19 (7)
Prostate-specific antigen, ng/mL†	1.37 (0.94)	1.72 (1.1)

Values expressed as mean (SD).

* Dementia Rating Scale total score: 144 possible points, with 138–144 in the normal range.
† Normal range, 0–4 ng/mL.

Primary outcome measure. The primary outcome measure was a battery of cognitive tests assessing spatial ability and memory (spatial and verbal). These cognitive domains were expected to change in response to androgen manipulation. Secondary outcome measures of language and attention were included in the comprehensive battery as control measures, which were not expected to change in response to hormone administration. Inclusion of both test measures for cognitive domains expected to change and control measures for cognitive domains not expected to change were necessary to assess selective effects of hormones. Comparable versions of each test were administered at each time point. Psychometrists and participants were blind to the treatment condition.

Spatial memory measures. *Route Test.* This test measured the ability to navigate a short route within a room and is based on previous route learning studies.[13,14] The task used a 6×24-foot piece of black flooring on which a diamond pattern was placed with bright yellow tape. In the non-landmark version, the examiner created a particular route on the grid by using a bright red ribbon. The subject was asked to walk the route as shown. The ribbon was removed, and the subject was asked to immediately retrace the route without the ribbon. Three trials were administered followed by three trials of a new route, using pictures placed on the floor as landmarks. Then a delayed recall of both routes was assessed after 20 minutes. Perfor-

mance was assessed by calculating the number of correct sequential units summed across all trials. This test has been shown to have good reliability in both young and old populations and validity when compared with other route tests and in comparison of brain-injured with controls.[13]

Spatial Array Learning Test (SALT). This measure of spatial memory was adapted from the Visual Spatial Learning Test.[15] Participants were shown seven unique figures in a particular pattern placed on a grid. The subject was allowed to look at the designs and placement briefly and then asked to choose the correct designs from eight distractor designs and to place them in the correct position on the grid. This procedure was repeated for a series of five trials and after a 20-minute delay. Number of correct tokens placed in the correct location was recorded for each trial and summed across trials. Reliability and validity assessed in an older adult population is very good.[15]

Oculomotor delayed response (OMDR). This measure of memory for spatial location involves immediate or working memory at short delays (a time-limited form of memory mediated by prefrontal cortex in which information is held actively on-line) and declarative memory at delays of 30 seconds.[16,17] Participants were required to focus on a central fixation point on a computer touch screen. While participants remain fixated, a cue appeared in one of 32 possible locations at a 4.5-inch radius from a central fixation point for 100 ms, a duration too brief for any eye movements to occur. The screen cleared and a series of single digits replaced the central fixation point. Participants were required to read each digit aloud during a delay period of 10 seconds or 30 seconds. After the delay, participants touched the place on the computer screen where they remembered seeing the cue. Responses were recorded in X and Y coordinates, and degree of displacement was calculated as the distance in millimeters between the actual target and the participant's response. Test–retest reliability in humans has not been established, although validity in comparison with other spatial memory measures is good.[16,17]

Verbal memory measures. *Proactive interference (PI).* Participants listened to a list of 10 words from the same semantic category (e.g., articles of clothing), and then recalled as many of these words as possible. The task was adapted from a previous task.[18] The procedure was repeated for a total of four trials, each containing different words drawn from the same semantic category. For the fifth trial, 10 words from a new semantic category (e.g., types of furniture) were read, and participants were asked to recall these words. The total number of words recalled correctly on each trial was recorded. Normal adults recall progressively fewer words across trials 2 through 4 because of the build-up of interference from the semantically similar preceding items. Reliability of the test is generally good, including validity studies conducted with brain-damaged patients and controls.[19]

Story Recall. The Story Recall task was modeled on the Wechsler Memory Scale–Revised (WMS-R) and measured memory for aurally presented contextual material. Participants listened to two brief narratives, each containing 25 informational bits, and were asked to recall as much as possible immediately after hearing each story and after a 20-minute delay.[20] Total number of words recalled from both stories was summed for immediate and delayed recall. Delayed recall was divided by immediate recall to obtain a savings score or a percentage of information retained. Reliability and validity of WMS-R, WMS-III Logical Memory, and this modified version are very good.[20–24]

Spatial ability measure. *Block Design.* This test is based on the Wechsler Adult Intelligence Scale–Revised (WAIS-R), Block Design subtest and measures participants' ability to analyze and construct abstract figures from their component parts.[25] The subject is shown individual red and white designs on paper and asked to construct the design by using nine three-dimensional blocks with red and white sides. Time to completion is recorded for each design, with an upper limit of 3 minutes per design. There are nine total designs, with three designs per difficulty level (easy, moderate. and hard). Total time for hard designs was summed for analysis. Reliability and validity of the WAIS-III Block Design subtest are well established.[26]

Selective attention measure. *Stroop Color Word Interference Task.* This task was based on the original version and used three trials, for which total reading time and errors were recorded.[27] The first condition (word reading) required participants to read 100 color words (*red, green, blue*), presented in rows on a sheet of paper, as quickly as possible. The second condition (color naming) required participants to name the color of 100 colored blocks presented in rows on a sheet of paper. In the third condition (color word interference), stimuli consist of color names that are printed in discordant colors (e.g., the word *blue* printed in green letters). Participants were asked to name the ink color of the printed words and are thus required to inhibit the reading of the words themselves. Because only the interference trial is a measure of divided attention, total time for trial three was used as the dependent measure. The test has shown good reliability and validity when examined in closed head–injured individuals compared with controls.[19]

Verbal ability measure. *Verbal fluency.* The verbal fluency measure is based on the Controlled Oral Word Association Test (COWAT) and is sensitive to verbal dysfunction and frontal lobe functioning. Participants were asked to verbally generate as many words beginning with a particular letter (e.g., *P*) within a 60-second period.[28] Two trials were administered with two different letters. The total number of words generated was recorded for each letter and summed. Test–retest reliability in an older population is good (0.70; 0.71) and validity studies using factor analysis suggest that it loads on a factor of abstract mental operation and language.[19,29]

Hormone assays. Blood samples were drawn weekly at the time of clinic visit, approximately 7 days after injection. For cognitive testing appointments, hormone levels were drawn immediately after cognitive testing that occurred 24 to 48 hours after injection. Samples were kept frozen in a −56 °C freezer until the completion of the study, when all samples were included in the assays. Serum estradiol and total testosterone levels were analyzed with RIA according to standard procedures for each commercial kit. Testosterone assays were run with a DELFIA testosterone kit (Wallac OY, Gaithersburg, MD), with sensitivity of 0.5 nmol/L and 4.5% intra-assay variability. 17β-Estradiol was measured with the ImmunuChem double-antibody kit (ICN Biomedicals, Inc., Costa Mesa, CA), with

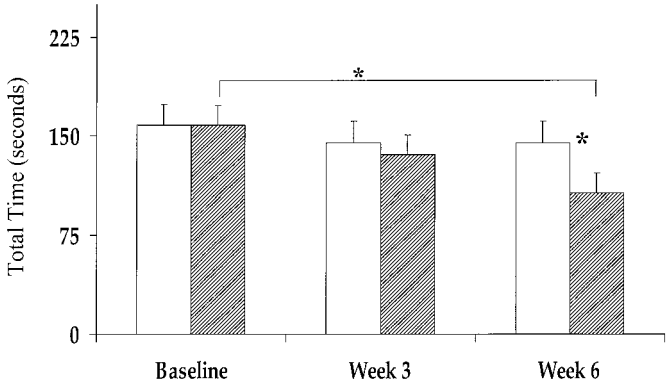

*Figure 2. Mean total time in seconds to complete Block Design at baseline and weeks 3 and 6 of treatment. Striped or gray bars represent the testosterone-treated group, and open bars represent the placebo group. SE bars represent standard error of measurement. The asterisk located between bars indicates a significant difference between groups (placebo and testosterone treated). The asterisk and line above bars indicates a significant change from baseline in the testosterone-treated group at week 6 (*p < 0.05).*

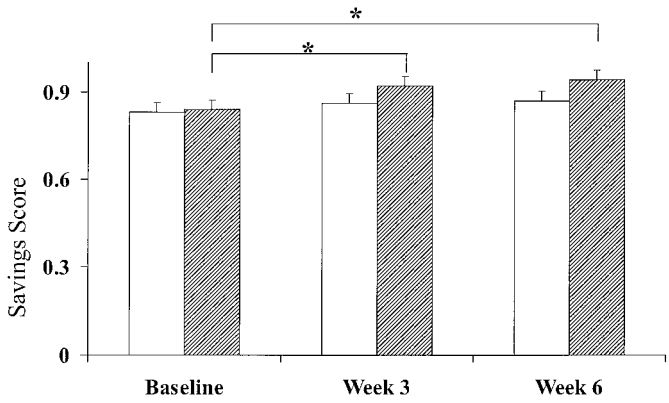

*Figure 3. Mean savings score on Story Recall, a measure of verbal memory, at baseline and weeks 3 and 6 of treatment. Savings score is the number of exact words recalled from the story after a 20-minute delay divided by the number of words recalled immediately after hearing the story. Therefore, savings score represents the percentage of information recalled. Striped or gray bars represent the testosterone-treated group, and open bars represent the placebo group. SE bars represent standard error of measurement. Asterisks and lines above bars indicate significant changes from baseline in the testosterone-treated group at weeks 3 and 6, as indicated by the lines (*p < 0.05).*

36.7 pmol/L sensitivity and 6.75 intra-assay variability. Samples from each participant were run in duplicate in the same assay to avoid interassay variability. Hormone levels reported were the average of the duplicate samples.

Statistical analyses. The primary outcome measure was to assess differences between the treatment group and the placebo group during two on treatment time-points on a battery of neuropsychological tests measuring memory and spatial ability. A repeated measures, multivariate analysis of variance (MANOVA) was used with group as the independent factor (treatment vs placebo) and weeks (baseline, weeks 3 and 6) as the repeated factor, spatial ability (Block Design), spatial memory (SALT, Route Test, OMDR), and verbal memory (PI and Story Recall) were dependent measures. Planned comparisons of on treatment time points (weeks 3 and 6) compared with baseline were performed, and post hoc comparisons were subjected to Bonferroni correction. Secondary outcome measures included measures of cognitive domains that were not expected to change from androgen manipulation (verbal fluency, Stroop test). A repeated measures MANOVA was used with group as the independent factor (treatment vs placebo) and weeks (baseline, weeks 3 and 6) as the repeated factor, and attention (Stroop) or language (verbal fluency) as dependent measures. Circulating plasma hormone levels were assessed using a repeated measures MANOVA, with group as the independent factor (treatment vs placebo), weeks (baseline, weeks 3 and 6) as the repeated factor, and total testosterone and total estradiol as the dependent measures. Information from the washout visit (week 12) was not included in the analysis for two reasons: 1) the primary study question was to assess whether manipulation of circulating hormone levels in healthy older men resulted in subsequent changes in cognitive functioning, which can be assessed by evaluating baseline and on treatment time points; and 2) recent evidence indicates that changes in cognition from hormone manipulation do not return to baseline levels but remain or continue to change in the same direction after cessation of hormone treatment.[30]

Results. *Primary outcome measure.* Repeated-measures MANOVA, with group as the independent factor (treatment vs placebo) and weeks (baseline, weeks 3 and 6) as the repeated factor, and spatial ability (Block Design), spatial memory (SALT, Route Test, OMDR), and verbal memory (proactive interference and Story Recall) as dependent measures indicated a significant omnibus multivariate effect of weeks ($F[7,12] = 4.78$), but no significant interaction effect. Significant change across weeks occurred in the treatment

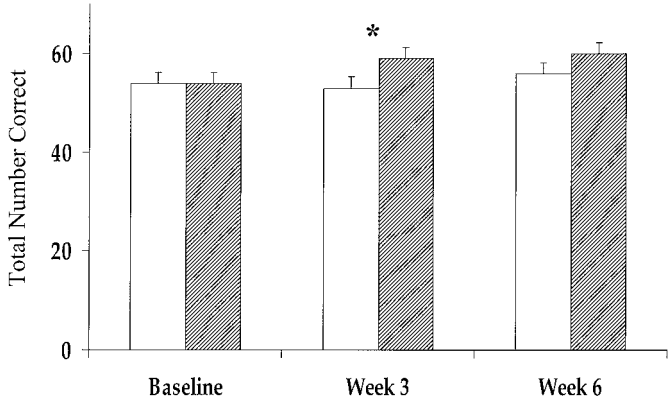

*Figure 4. Mean total points correct on the Route Test, a measure of spatial and navigational memory at baseline and weeks 3 and 6 of treatment. Striped or gray bars represent the testosterone-treated group, and open bars represent the placebo group. A total of 72 points were possible. SE bars represent standard error of measurement. Asterisks located between bars indicate a significant difference between groups (placebo and testosterone treated) (*p < 0.05).*

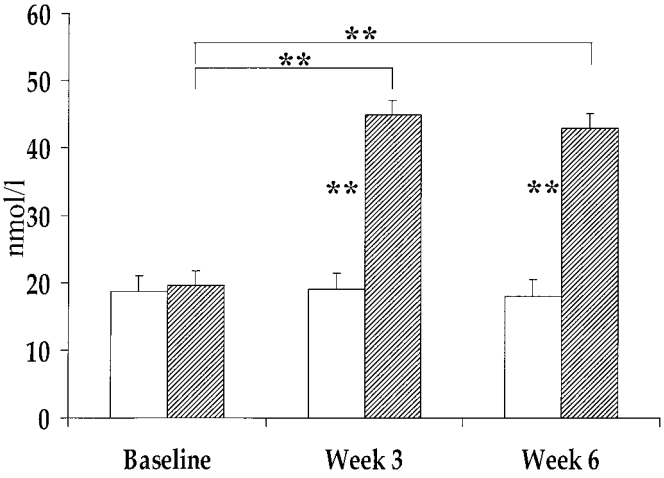

*Figure 5. Mean serum total testosterone levels (nmol/L) in treatment and placebo groups at baseline and weeks 3 and 6 of treatment. Striped or gray bars represent the testosterone-treated group, and open bars represent the placebo group. SE bars represent standard error of measurement. Asterisks located between bars indicate significant differences between groups (placebo and testosterone treated). Asterisks and lines above bars indicate significant changes in the testosterone-treated group at weeks 3 and 6 compared with baseline, as indicated by the lines (**p < 0.01, *p < 0.05). Significant interaction effect was also evident.*

group only (F[7,12] = 4.189). Planned comparisons between on treatment (weeks 3 and 6) and baseline (week 0) for each test revealed significant changes for Block Design at week 6 (F[1,7] = 6.14); Story Recall at week 6 (F[1,7] = 6.62), and

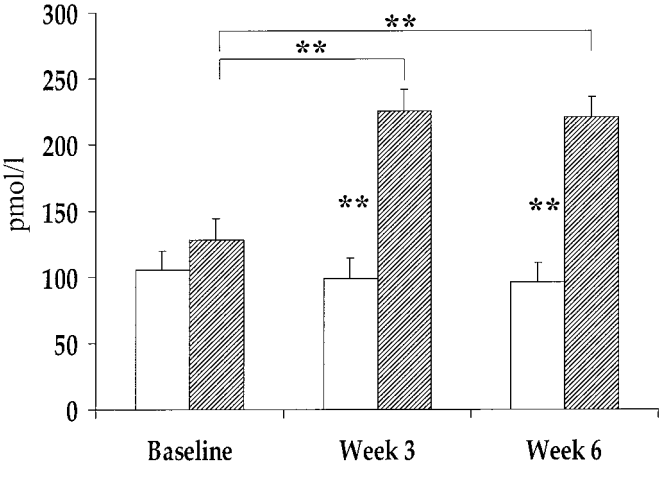

*Figure 6. Mean serum total estradiol levels (pmol/L) in treatment and placebo groups at baseline and weeks 3 and 6 of treatment. Striped or gray bars represent the testosterone-treated group, and open bars represent the placebo group. SE bars represent standard error of measurement. Asterisks located between bars indicate significant differences between groups (placebo and testosterone treated). Asterisks and lines above bars indicate significant changes in the testosterone-treated group at weeks 3 and 6 compared with baseline, as indicated by the lines (**p < 0.01, *p < 0.05). Significant interaction effect was also evident.*

OMDR at week 6 (F[1,7] = 17.4). Subsequent pairwise comparisons revealed significant improvement in the treatment group for Block Design at week 6 (figure 2); Story Recall at week 6 (figure 3), and both treatment and placebo group at week 6 for OMDR. Univariate, between-group comparisons revealed significant difference between groups on the Route Test at week 3 (F[1,18] = 5.22), with treatment group outperforming the placebo group (figure 4); and significant difference between groups on Block Design at week 6 (F[1,18] = 5.88), with treatment group outperforming placebo group (see figure 2). No significant differences between groups were noted at baseline. An examination of means for the treatment group at washout visit showed no large increases or decreases in performance for SALT, PI, verbal fluency, or Stroop. Block Design and Story Recall performance returned to baseline levels, and Route Test performance remained improved at on treatment levels.

Secondary outcome measures. Repeated-measures MANOVA, with group as the independent factor (treatment vs placebo) and weeks (baseline, weeks 3 and 6) as the repeated factor, and attention (Stroop) or language (verbal fluency) as dependent measures, indicated no significant omnibus multivariate effect of group, weeks, or group-by-week interaction. Planned comparisons and pairwise comparisons were nonsignificant.

Hormone measures. Repeated-measures MANOVA was performed with group as the independent factor (treatment vs placebo), weeks (baseline, weeks 3 and 6) as the repeated factor, and total circulating testosterone and estradiol levels as the dependent measure. Omnibus multivariate comparison of group (F[2,20] = 22.29), weeks (F[4,18] = 38.4), and week-by-group interaction (F[4,18] = 32.9) were all significant. Multivariate examination of weeks showed significant differences between baseline and weeks 3 (F[2,20] = 35.6) and 6 (F[2,20] = 24.1), with significant change across weeks evident for treatment group only (F[4,18] = 68.2). Subsequent pairwise comparisons indicated significant differences between baseline and weeks 3 and 6 for treatment group only for both testosterone and estradiol. Univariate comparisons of group showed significant differences between groups at weeks 3 (F[1,21] = 74.8) and 6 (F[1,21] = 50.1) for testosterone, and at weeks 3 (F[1,21] = 18.3) and 6 (F[1,21] = 17.1) for estradiol (figures 5 and 6). No significant differences between groups were evident for baseline measures. Testosterone levels were raised from an average of 20 ± 11 nmol/L in the treatment group to the range of healthy young men (44 ± 5 nmol/L). PSA levels significantly increased at week 4 compared with baseline in the treatment group (F[2,21] = 5.54). However, as expected, PSA levels returned to baseline at washout visit, with no significant differences between screen and washout. One participant with a near-borderline PSA at the start of the study (3.5 ng/mL) had an elevated PSA at week 4 (4.1 ng/mL) and was subsequently discontinued from the study.

To further examine the relationship between hormone levels and cognitive function, Pearson correlation coefficients were calculated between cognitive measures and testosterone and estradiol levels taken at the time of testing. Significant correlations were observed between story recall savings score and estradiol and testosterone levels at week 6 of treatment. Selective attention as measured by the Stroop test showed a significant positive correlation

Table 2 *Pearson product-moment correlations between cognitive measures and serum testosterone (T) and estradiol (E_2) levels taken at the time of testing*

	Week 3		Week 6	
Cognitive domain/test	T, nmol/L	E_2, pmol/L	T, nmol/L	E_2, pmol/L
Spatial memory				
Route test	0.278	0.126	0.234	0.024
OMDR	0.057	−0.030	−0.184	−0.218
Spatial Array Learning Test	−0.131	−0.345	0.018	−0.118
Verbal memory				
Story Recall	0.204	0.246	0.454*	0.596†
Proactive interference	0.075	−0.208	0.161	−0.115
Constructional ability				
Block Design	−0.090	0.114	−0.389	−0.164
Language				
Verbal fluency	−0.137	−0.420*	0.233	−0.055
Attention				
Stroop	−0.006	0.202	0.060	0.468*

* $p < 0.05$.
† $p < 0.01$.

OMDR = oculomotor delayed response.

with estradiol levels at week 6, and verbal fluency showed a significant negative correlation with estradiol levels at week 3 of treatment (table 2). Because of multiple individual correlation calculations, correlation results must be interpreted with caution.

Discussion. Our results indicate that healthy older men show improvements in spatial and verbal memory and spatial abilities in response to short-term testosterone treatment. Although improvements were not evident for all measures, our results are consistent with previous findings of improved spatial abilities and spatial memory in response to testosterone administration.[9,31] These improvements were not evident for all cognitive domains, such as selective attention or language, suggesting that increases in serum testosterone or estradiol have selective or specific effects on memory and spatial abilities. The improvement in verbal memory may be considered clinically significant, because a 4 to 8% savings score increase represents a 1 to 2 standard deviation improvement as measured by a widely used clinical measure of verbal memory.[21]

Improved spatial memory was not found for all spatial memory measures. Significantly improved spatial memory was observed on the Route Test but not for spatial memory as measured by the SALT. Although modest improvements were noted on the OMDR, these were observed for both the treatment group and the placebo group. Therefore, improvement on the OMDR cannot be considered a significant improvement beyond an observed practice effect. It is possible that task demands unique to each test may be differentially affected by changes in hormones. For example, the Route Test requires a three-dimensional recall of spatial information (i.e., a walking route) in the absence of verbal information. In contrast, success on the SALT can be achieved with a nonspatial (i.e., verbal) strategy or a combination of these. A verbal strategy is difficult and unlikely to be used while navigating and recalling a walking route. Consistent with these results, studies of exogenous testosterone administration have shown improvements in spatial memory and spatial abilities, and a lack of improvement for visual memory tasks.[7,9,31] These results provide further evidence for a specific role of testosterone on spatial memory and spatial abilities. Several lines of evidence suggest that testosterone has a selective effect on spatial abilities. Male rats compared with female rats show larger and more asymmetric cell layers in the hippocampus, which underlies spatial memory and navigation.[32] Neonatal exposure of female rats to testosterone results in a more masculine (larger more asymmetrical) hippocampus with regard to size.[33] These organizational changes likely reflect differential effects of gonadal hormones on place and landmark systems in the hippocampus and can be reversed with early hormonal manipulation.[34] In humans, men tend to use a more spatial navigational strategy (e.g., geometric cues) when solving maze tasks, whereas women tend to use visual detail cues (e.g., landmarks).[16,35]

Changes in verbal memory in response to testosterone administration have not been previously reported. One possible explanation for these intriguing results is that they may be attributable to the rise in

estradiol levels occurring in response to testosterone administration, because the estradiol levels showed a significant positive correlation with the savings score. Significant improvements in verbal recall have been found in healthy older women and patients with AD receiving estrogen replacement treatment.[36,37] In men, exogenous increases in estrogen have been shown to improve verbal memory for a paired associate learning task, a task that generally favors women.[38] In the current study, improved verbal recall was noted for savings score on the story recall test, but not for verbal memory as measured by the PI test. This may be because of attention and executive function demands inherent in the PI test, because no improvements in the treatment group were noted for the Stroop test, a measure of selective attention, and a positive correlation was found between estradiol levels and poorer performance on the Stroop. It is also possible that increased testosterone affects storage or consolidation of verbal information as opposed to retrieval, which would result in a relative greater improvement on story recall compared with the PI test.

Androgen effects on memory may be mediated through several mechanisms. First, the hippocampus, which underlies spatial abilities and declarative memory, contains both testosterone receptors and estrogen receptors.[32,39] Therefore, testosterone may have direct effects on the hippocampus through the androgen receptor as well as indirect effects from aromatization to estradiol interacting with estradiol receptors.[40] Both types of estrogen receptors, ERα and ERβ, have been found in the rat hippocampus, and recent in vivo estrogen-binding studies suggest that estrogen is likely involved in cognition and neuroprotection in the hippocampus and basal forebrain.[41,42] Behavioral effects of ERα knockout mice also suggest that estrogen has effects on learning and memory, although the relative importance of ERα vs ERβ receptors is yet unknown.[43]

Second, there is some evidence that gonadal steroids may act through nongenomic mechanisms. Gonadal steroids have been shown to have rapid effects on the excitability of hippocampal pyramidal cells,[44] and recent studies indicate that gonadal steroids can act via the cell surface through interaction with specific neurotransmitter receptors.[45-47] Overall, it is likely that gonadal steroids act through both genomic and nongenomic mechanisms, creating a unique intracellular cross-talk.[47,48]

Although these results will need to be further replicated, they provide a rationale for further studies examining the effects of gonadal steroids on mood and cognition, particularly in healthy older men and women who experience natural declines in hormones from aging. It is unclear whether the observed improvements in verbal and spatial memory and spatial abilities are attributable to testosterone or estradiol from aromatization or both. Therefore, further studies will need to examine the relative contribution of testosterone vs estradiol on cognition in men. Finally, our findings of improved verbal and spatial memory in response to testosterone also provide a strong rationale for examining the effects of gonadal steroids in patients with AD who experience significant impairments in memory. However, a recent, double-blind, placebo controlled, multicenter trial of hormone replacement therapy (HRT) in women patients with AD with hysterectomies did not show cognitive benefits for the HRT group as compared with placebo.[49] The results of this study are difficult to interpret, because only patients with AD with hysterectomies or with both hysterectomy and oophorectomy were included, and Premarin, the form of HRT used, contains many different forms of estrogen and may not be the most potent form of estrogen available. In contrast, a recent, double-blind, placebo-controlled study using 17β-estradiol, a potent form of estrogen, found cognitive improvements in verbal memory and attention in women with AD.[36] However, the treatment period was brief, and cognitive improvements were observed in verbal memory and attention as measured by objective and sensitive psychometric tests. Clearly, further research is needed to better elucidate the relationship between hormones and cognition.

Acknowledgment

The authors thank Karla Grimwood, Cassin Lofgreen, Lyndel Cubberley, Jody Lowry, Jane Hammett, Allison Weston, Maria Gonzales, Kristina Purganan, and Colby Wait for technical assistance.

References

1. Tenover JS, Matsumoto AM, Plymate SR, et al. The effects of aging in normal men on bioavailable testosterone and luteinizing hormone secretion: response to clomiphene citrate. J Clin Endocrinol Metab 1987;65:1118–1126.
2. Tenover J. Effects of testosterone supplementation in the aging male. J Clin Endocrinol Metab 1992;75:1092–1098.
3. Swerdloff RS, Wang C. Androgens and aging men. Exp Gerontol 1993;28:435–446.
4. Tenover JS. Androgen administration to aging men. Endocrinol Metab Clin North Am 1994;23:877–892.
5. Flood JF, Farr SA, Kaiser FE, et al. Age-related decrease of plasma testosterone in SAMP8 mice: replacement improves age-related impairment of learning and memory. Physiol Behav 1995;57:669–673.
6. Morely JE, Kaiser F, Raum WJ, et al. Potentially predictive and manipulable blood serum correlates of aging in the healthy human male: progressive decreases in bioavailable testosterone, dehydroepiandrosterone sulfate, and the ratio of insulin-like growth factor to 1 growth hormone. Proc Natl Acad Sci 1997;94:7537–7542.
7. Sih R, Morley JE, Kaiser FE, et al. Testosterone replacement in older hypogonadal men: a 12 month randomized controlled trial. J Clin Endocrinol Metab 1997;82:1661–1667.
8. Wolf O, Pruet R, Hellhammer DH, et al. Testosterone and cognition in elderly men: a single testosterone injection blocks the practice effect in verbal fluency, but has no effect on spatial or verbal memory. Biol Psychiatry 2000;47:650–654.
9. Janowsky JS, Chavez B, Orowoll E. Sex steroids modify working memory. J Cogn Neurosci 2000;12:407–414.
10. Matsumoto AM. Hormonal therapy of male hypogonadism. Endocrinol Metab Clin North Am 1994;23:857–875.
11. Folstein MF, Folstein MF, McHugh PR. Mini-Mental state: a practical method for grading the cognitive state of patients for the clinician. J Psychiatr Res 1975;12:189–198.

12. Mattis S. Dementia rating scale. Odessa, FL: Psychological Assessment Resources, Inc., 1988.
13. Barrash J, Tranel D. Neuropsychological correlates of route learning. J Int Neuropsychol Soc 1996;2:69.
14. Cherrier MM, Mendez MF. Performance of Alzheimer's disease patients and healthy controls on an ecologically valid route learning test. J Int Neuropsychol Soc 1997;3:66.
15. Malec JF, Ivnik RJ, Smith GE, et al. Visual Spatial Learning Test: normative data and further validation. Psychological Assess 1992;4:433–441.
16. Wilson FA, Riches IP, Brown MW. Hippocampus and medial temporal cortex: neuronal activity related to behavioral responses during the performance of many tasks by primates. Behav Brain Res 1990;40:7–28.
17. Mishkin M, Murray EA. Stimulus recognition. Curr Opin Neurobiol 1994;4:200–206.
18. Moscovitch M. Cognitive resources and dual-task interference effects on retrieval in normal people: the role of the frontal lobes and medial temporal cortex. Neuropsychology 1994;8: 524–534.
19. Lezak MD. Neuropsychological assessment. 3rd ed. New York, NY: Oxford University Press; 1995.
20. Wechsler D. Wechsler Memory Scale–Revised. San Antonio, TX: The Psychological Corporation, 1987.
21. Wechsler D. Wechsler Memory Scale III. San Antonio, TX: The Psychological Corporation, 1997.
22. Craft S, Murphy C, Wemstrom J. Glucose effects on memory and complex non-memory measures: the influence of age, sex, and glucoregulatory response. Psychobiology 1994;22:95–105.
23. Craft S, Dagogo-Jack E, Wiethop CM, et al. The effects of hyperglycemia on memory and hormone levels in dementia of the Alzheimer's type: a longitudinal study. Behav Neurosci 1993;107:926–941.
24. Craft S, Zallen G, Baker LD. Glucose and memory in mild senile dementia of the Alzheimer type. J Clin Exp Neuropsychol 1992;14:253–267.
25. Wechsler D. Wechsler Adult Intelligence Scale–Revised. San Antonio, TX: The Psychological Corporation, 1981.
26. Wechsler D. WAIS-III administration and scoring manual. San Antonio, TX: The Psychological Corporation, 1997.
27. Stroop JR. Studies of interference in serial verbal reactions. J Exp Psychol 1935;18:643–662.
28. Spreen O, Strauss E. A compendium of neuropsychological tests. New York: Oxford University Press; 1991.
29. Snow WG, Tierney MC, Zorzitto ML. One-year test re-test reliability of selected neuropsychological tests in older adults. J Clin Exp Neuropsychol 1988;10:60.
30. Slabbekoorn D, van Goozen SH, Megens J, et al. Activating effects of cross-sex hormones on cognitive functioning: a study of short-term and long-term hormone effects in transsexuals. Psychoneuroendocrinology 1999;24:423–447.
31. Janowsky JS, Oviatt SK, Orwoll ES. Testosterone influences spatial cognition in older men. Behav Neurosci 1994;108:325–332.
32. Roof RL, Havens MD. Testosterone improves maze performance and induces development of a male hippocampus in females. Brain Res 1992;572:310–313.
33. Roof RL. The dentate gyrus is sexually dimorphic in prepubescent rats: testosterone plays a significant role. Brain Res 1993;610:148–151.
34. Williams CL, Barnett AM, Meck WH. Organizational effects of early gonadal secretions on sexual differentiation in spatial memory. Behav Neurosci 1990;104:84–97.
35. Galea LA, Kimura D. Sex differences in route learning. Personality and Individual Differences 1993;14:53–65.
36. Asthana S, Craft S, Baker LD, et al. Cognitive and neuroendocrine response to transdermal estrogen in postmenopausal women with Alzheimer's disease: results of a placebo-controlled, double-blind, pilot study. Psychoneuroendocrinology 1999;24:657–677.
37. Hogervorst E, Boshuisen M, Riedel W, et al. Richter Award. The effect of hormone replacement therapy on cognitive function in elderly women. Psychoneuroendocrinology 1999;1998: 24:43–68.
38. Miles C, Green R, Sanders G, et al. Estrogen and memory in a transsexual population. Horm Behav 1998;34:199–208.
39. Luine VN. Steroid hormone influences on spatial memory. Ann NY Acad Sci 1994;743:201–211.
40. Naftolin F, Ryan KJ, Davies IJ, et al. The formation of estrogens by central neuroendocrine tissues. Recent Prog Horm Res 1975;31:295–319.
41. Shughrue PJ, Scrimo PJ, Merchenthaler I. Estrogen binding and estrogen receptor characterization (ERalpha and ERbeta) in the cholinergic neurons of the rat basal forebrain. Neuroscience 2000;96:41–49.
42. Shughrue PJ, Merchenthaler I. Estrogen is more than just a "sex hormone": novel sites for estrogen action in the hippocampus and cerebral cortex. Front Neuroendocrinol 2000; 21:95–101.
43. Rissman EF, Wersinger SR, Fugger HN, et al. Sex with knockout models: behavioral studies of estrogen receptor alpha. Brain Res 1999;835:80–90.
44. Teyler T, Vardaris R, Lewis D, et al. Gonadal steroids: effects on excitability of hippocampal pyramidal cells. Science 1980; 209:1017–1019.
45. Schmidt BM, Gerdes D, Feuring M, et al. Rapid, nongenomic steroid actions: a new age? Front Neuroendocrinol 2000;21:57–94.
46. Compagnone NA, Mellon SH. Neurosteroids: biosynthesis and function of these novel neuromodulators. Front Neuroendocrinol 2000;21:1–56.
47. Rupprecht R, Holsboer F. Neuroactive steroids: mechanisms of action and neuropsychopharmacological perspectives. Trends Neurosci 1999;22:410–415.
48. Rupprecht R. The Neuropsychopharmacological potential of neuroactive steroids. J Psychiatr Res 1997;31:297–314.
49. Mulnard RA, Cotman CW, Kawas C, et al. Estrogen replacement therapy for treatment of mild to moderate Alzheimer disease: a randomized controlled trial. Alzheimer's Disease Cooperative Study [see comments]. JAMA 2000;283:1007–1015.