Distinct cognitive effects of estrogen and progesterone in menopausal women

Alison Berent-Spillsona,*, Emily Bricenob, Alana Pinskyc, Angela Simmend, Carol C. Persadb, Jon-Kar Zubietaa, and Yolanda R. Smithd,∗

aUniversity of Michigan, Psychiatry Department. MBNI, 205 Zina Pitcher Place, Ann Arbor, MI 48109, USA
bUniversity of Michigan, Psychiatry Department, Neuropsychology Division, 2101 Commonwealth Blvd, Suite C, Ann Arbor, MI 48105, USA
cUniversity of Michigan Medical School, 1301 Catherine, Ann Arbor, MI 48109, USA
dUniversity of Michigan, Obstetrics and Gynecology Department, L4000 Womens SPC, 1500 E. Medical Center Dr, Ann Arbor, MI 48109, USA

Abstract

The effects of postmenopausal hormone treatment on cognitive outcomes are inconsistent in the literature. Emerging evidence suggests that cognitive effects are influenced by specific hormone formulations, and that progesterone is more likely to be associated with positive outcomes than synthetic progestin. There are very few studies of unopposed progesterone in postmenopausal women, and none that use functional neuroimaging, a sensitive measure of neurobiological function.

In this study of 29 recently postmenopausal women, we used functional MRI and neuropsychological measures to separately assess the effects of estrogen or progesterone treatment on visual and verbal cognitive function. Women were randomized to receive 90 days of either estradiol or progesterone counterbalanced with placebo. After each treatment arm, women were given a battery of verbal and visual cognitive function and working memory tests, and underwent functional MRI including verbal processing and visual working memory tasks.

*Corresponding author: Alison Berent-Spillson, 1-734-615-4252.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

We have not published this work previously and it is not under consideration for publication elsewhere. This publication as approved by all authors, and if accepted, it will not be published elsewhere in the same form without the written consent of the copyright-holder.

All authors have approved the final article, and include:
Alison Berent-Spillson: study design, data collection, preparation, analysis, interpretation, manuscript preparation
Emily Briceno: data analysis, interpretation, manuscript preparation
Alana Pinsky: data preparation, analysis, interpretation
Angela Simmen: data interpretation, manuscript preparation
Carol Persad: study design, data interpretation
Jon-Kar Zubieta: study design, data interpretation
Yolanda Smith: study design, data interpretation, manuscript preparation
We found that both estradiol and progesterone were associated with changes in activation patterns during verbal processing. Compared to placebo, women receiving estradiol treatment had greater activation in the left prefrontal cortex, a region associated with verbal processing and encoding. Progesterone was associated with changes in regional brain activation patterns during a visual memory task, with greater activation in the left prefrontal cortex and right hippocampus compared to placebo. Both treatments were associated with a statistically nonsignificant increase in number of words remembered following the verbal task performed during the fMRI scanning session, while only progesterone was associated with improved neuropsychological measures of verbal working memory compared to placebo. These results point to potential cognitive benefits of both estrogen and progesterone.

Keywords
Progesterone; postmenopausal hormone treatment; cognition; menopause; functional MRI

1. Introduction
The cognitive effects of postmenopausal hormone treatment have remained controversial in recent years, and as a result, hormones are less commonly prescribed for complaints of memory loss. However, hormones are still commonly prescribed on and off-label for non-cognitive indications, and some hormone formulations may impart a cognitive benefit for some women. There is emerging evidence that progestins and progesterone do not equivalently influence neurobiological mechanisms of cognitive function, and that progesterone is likely to be more beneficial and carry fewer risks than its synthetic counterparts (Fischer et al., 2014; Jodhka et al., 2009; Maki, 2012; Singh and Su, 2013a). Despite this distinction, there is currently little evidence on the cognitive effects of progesterone in postmenopausal women.

Postmenopausal hormone use by women with intact uteri consists of a combined estrogen and progesterone regimen. Previously standard conjugated equine estrogen (CEE) + medroxyprogesterone acetate (MPA) has been associated with several health risks, such as increased risk of breast cancer and cardiovascular disease, and variable cognitive effects, including negative cognitive outcomes in women who began treatment well past menopause (Braden et al., 2011; Chlebowski et al., 2013; Coker et al., 2010; Resnick et al., 2006; Rossouw et al., 2002; Shumaker et al., 2003). Synthetic progestins have also been associated with reductions in verbal ability in all postmenopausal age groups (Coker et al., 2010; Maki et al., 2007; Resnick et al., 2004; Resnick et al., 2006), however verbal impact appears to be heavily influenced by progestin formulation. Progestins with less anti-estrogenic effects have been shown to have neutral or positive impact on verbal benefits conferred by estradiol treatment (Spark and Willis, 2012).

Recent studies of combined hormone formulations suggest that progesterone carries fewer somatic risks than synthetic progestins, and is not associated with increased risk for adverse cardiovascular outcomes, venous thromboembolism, or breast cancer (L’Hermite, 2013; Simon, 2012). Progesterone is also documented to have neuroprotective effects (De Nicola
et al., 2013; Deutsch et al., 2013; Melcangi et al., 2014; Singh, 2006; Singh and Su, 2013b). Natural progesterone is associated with more positive and fewer negative cognitive outcomes than synthetic progestins, particularly MPA. In fact some evidence suggests that type of progestogenic compound is more critical than type of estrogen in determining cognitive impact (Maki, 2012; Stanczyk et al., 2013; Warren et al., 2014).

Despite the significant distinctions between progesterone and synthetic progestins, there are very few studies that isolate the effects of progesterone on cognitive outcomes in postmenopausal women, and none using neuroimaging measures of cognitive processing. A single study of cognitive effects of postmenopausal progesterone administration found no effect on verbal or executive function performance, but did not test visual cognition (Schussler et al., 2008; Spark and Willis, 2012). The majority of evidence for cognitive effects of natural progesterone comes from studies of postmenopausal women using combined hormone formulations (estrogen + progestin versus progesterone), studies of circulating progesterone levels or fluctuations in endogenous progesterone levels across a menstrual cycle, and progesterone administration in animals. Evidence from these studies suggest an effect of progesterone in cognitive processing brain regions, distinct from the effects of estrogen, and indicate the need for closer investigation of progesterone effects in postmenopausal women (Acosta et al., 2013; Chisholm and Juraska, 2012; Drake et al., 2000; Farage et al., 2008; Gibbs, 2000; Nilsen and Brinton, 2002b; Rapp et al., 2003; Stanczyk et al., 2013).

While cognitive outcomes have been variable across studies of postmenopausal hormone use, in general there have been more positive cognitive effects associated with unopposed estrogen than with combined estrogen/progestin regimens, with progestin appearing to mitigate estrogenic effects in some cases (Rice et al., 2000; Sherwin and Grigorova, 2011; Silverman et al., 2011). This is particularly true for the verbal cognitive domain, which appears to be more influenced by hormone treatment than visual cognition (Carlson and Sherwin, 1999; Jacobs et al., 1998; Krug et al., 2003; Maki, 2005; Resnick et al., 2006; Shaywitz et al., 2003; Sherwin, 2012). It is possible that synthetic progestins reduce some neuroprotective effects of estrogen, as has been demonstrated in cell culture models, and these studies also suggest that natural progesterone may increase estrogenic neuroprotections (Nilsen and Brinton, 2002a, b).

In contrast to variable behavioral outcomes, a growing collection of imaging studies of menopausal estrogen use have almost uniformly found differing brain activation patterns between hormone users and non-users during both verbal and visual cognitive tasks (Berent-Spillson et al., 2010; Dumas et al., 2010; Joffe et al., 2006; Persad et al., 2009; Shaywitz et al., 1999; Smith et al., 2006). Despite significant differences in study design, including differences in length of treatment, hormone formulations, and cognitive tasks, functional MRI (fMRI) of estrogen use has routinely shown increased activation during cognitive processing in frontal and cingulate cortical regions of the working memory circuitry (Dumas et al., 2010; Joffe et al., 2006; Shaywitz et al., 1999). Results from fMRI studies of combined estrogen + progestin have been less unequivocal, with similar but less distinct differences in activation patterns compared to non-users during verbal and visual cognitive tasks (Persad et al., 2009; Smith et al., 2006). Despite the distinct differences in activation of
cognitive association and working memory circuitry between hormone users and non-users, sample sizes of neuroimaging studies have largely precluded meaningful behavioral analyses. However increased activation was associated with better visual working memory task performance in long-term estrogen or estrogen + progestin users (Berent-Spillson et al., 2010), and hormone use was associated with differential hippocampal activation and better verbal working memory performance in early-initiation hormone users compared to non-users (Maki et al., 2011). PET studies of blood flow to cognitive processing circuitry have also indicated differences between hormone users and non-users, largely but not uniformly indicating increased blood flow to cognitive association regions in hormone users during verbal and visual tasks, and often correlated with better cognitive task performance (Maki and Resnick, 2000; Resnick et al., 1998). To date there have been no neuroimaging studies of cognitive effects of unopposed progestin or progesterone use during menopause, however a notable PET study of Lupron-induced ovarian suppression in young women found that estrogen and progesterone were each independently able to regulate cerebral blood flow and prefrontal cortical activation during a test of frontal executive function (Berman et al., 1997).

Because of the differential effects of hormone use on visual and verbal cognition, and because of the neuroprotective potential of progesterone, there is a need for further study of the effects of estrogen and progesterone independently on each of these cognitive domains. In the current randomized placebo-controlled pilot study of 29 recently postmenopausal women, we separately assessed the effects of estrogen or progesterone treatment on visual and verbal cognitive function, using both neuropsychological measures and functional MRI to examine neural pathways used while performing verbal processing and visual working memory cognitive tasks. We hypothesized that estrogen treatment would be associated with more robust changes in cognitive performance and regional brain activation compared to placebo than would progesterone treatment, particularly in the prefrontal cortex and hippocampus, and that progesterone might be associated with slightly worse verbal processing performance scores.

2. Methods

2.1 Subjects

Twenty-nine women (25 included in fMRI imaging analysis), aged 45 – 55 years were recruited from the community. Women were included with FSH values over 40 IU/L and serum estradiol less than 40 pg/ml, and between 6–38 months of amenorrhea, to include women close to their final menstrual period. Two women with hysterectomy and bilateral oophorectomy meeting all other criteria were included, one in each of the two hormone treatment groups. Women were excluded for left handedness, acute illness, neurological disease, current or past psychiatric illness, including depression, bipolar disorder, anxiety disorder, and schizophrenia, migraines, heart, thromboembolic, liver, or uncorrected thyroid disease, diabetes, porphyria, allergy to estradiol, progesterone, or lactose, contraindications to MRI (claustrophobia, internal metallic devices), smoking within 3 years, hormones within 3 months, history of substance abuse, head injury, or loss of consciousness, centrally acting medications, endometrial lining over 5mm, ovarian pathology, abnormal mammogram,
creatinine levels over 1.5 mg/dl, fasting cholesterol greater than 300 mg/dl, and fasting triglycerides greater than 300 mg/dl.

Of 30 women initially enrolled, one dropped out prior to treatment randomization. Of the 29 women who completed both treatment arms and clinical and behavioral assessments, two did not undergo fMRI scanning due to claustrophobia (both from the estrogen group), one did not undergo fMRI scanning after active treatment (estrogen), and one completed all scans but was excluded from fMRI data analysis due to incomplete image data files (from the progesterone group).

2.2 Study protocol

This was a double-blind placebo-controlled randomized pilot study of cognitive and neurobiological responses to short-term (12 weeks) oral estradiol (TEVA Pharmaceutical Industries Ltd, Petach Tikva, Israel; 1mg; n=13) or progesterone (Letco Medical, Decatur, AL; 200mg; n=16) treatment. Women underwent clinical evaluation including assessment of reproductive hormone levels, neuropsychological assessment of cognitive function, and fMRI to observe neural activation patterns during verbal and visual cognitive tasks. All procedures were approved by the University of Michigan Institutional Review Board, and written informed consent was obtained.

2.3 Clinical evaluations

Clinical evaluations were performed at the Michigan Clinical Research Unit (MCRU) at the University of Michigan, and included physical examination, and vaginal ultrasound to assess endometrial lining. Psychiatric interview and assessment were performed to rule out current psychiatric illness, and women were evaluated for depression using the Hamilton Depression Rating Scale (Hamilton, 1960). Pap smears were performed and mammograms scheduled if not completed within the prior year. Serum was collected before 11:00 AM after an 8-hour fast for hormone assessment and to confirm study eligibility. Estradiol concentrations were measured with a modified off-line ACS:180 E2-6 immunoassay (Bayer Diagnostics Corp, Norwood, MA). FSH concentrations were measured with a two-site chemiluminometric immunoassay using two monoclonal antibodies with specificity for intact FSH (Bayer Diagnostics Corp). Progesterone concentrations were measured using a competitive chemiluminescent immunoassay using two antibodies (Siemens Medical Solutions USA, Inc., Malvern, PA).

2.4 Drug treatment

Study drugs were administered orally once a day at bedtime for 90 days. Active and placebo treatment arms were counterbalanced between subjects, and separated by a 12 week washout period (Table 1). Following active treatment, women randomized to estradiol took 200mg progesterone and women randomized to progesterone took placebo for 10 days to slough the endometrial lining. Randomization and study drug dispensing were coordinated by the Investigational Drug Service at the University of Michigan.
2.5 Neuropsychological assessments

Neuropsychological measures were administered by a trained neuropsychology technician, and were selected to measure verbal and visual cognitive function. Practice effects were minimized by using multiple test forms counterbalanced between treatment arms, and by a practice test session prior to treatment randomization to reduce the impact of general testing experience. The Hopkins Verbal Learning Test-Revised (HVLT-R) (Brandt, 1991) and Brief Visuospatial Memory Test-Revised (BVMT-R) (Benedict R.H.B. Schretlen D. Groninger L, 1996) were used to evaluate verbal and visual learning and memory, respectively. Verbal and visual working memory were measured with the Wechsler Memory Scale, Third Edition (WMS-III) Digit Span, Letter-Number Sequencing, and Spatial Span subtests (Wechsler, 1997) and the Brown Peterson paradigm (Strauss, 2006).

Delayed recall scores from the HVLT-R and the BVMT-R were used as outcome measures of verbal and visual delayed memory. A composite verbal working memory score was calculated from individual raw scores from the WMS-III Digit Span and Letter-Number Sequencing and Brown Peterson tasks. This composite was calculated by converting individual raw scores into a sample-based z-score (based upon the overall group mean and standard deviation), then averaging these z-scores to create the overall composite. To aid in interpretation of these values, z-scores were converted to standard scores (M=100, SD=15). Paired t-tests were used to compare scores after placebo and active treatment arms.

2.6 fMRI cognitive tasks

Episodic verbal memory processes (verbal encoding and storage) were assessed during the fMRI scanning session using the levels of processing paradigm, in which the women were asked to judge words based on physical characteristics (shallow processing) or on word meanings (deep processing) (Gabrieli et al., 1996). Sixteen lists of 12 words each were equated for letter length and frequency, with four counterbalanced lists presented in each of four scanning runs. Each list was preceded by 8 seconds of instructions, and each word presented for 1.5 seconds with a 1.5 second interstimulus interval. In the phonemic condition, subjects decided whether each word was presented in uppercase or lowercase letters (shallow processing), and in the semantic condition, whether each word represented an abstract or concrete concept (deep processing). Each list contained equal numbers of uppercase, lowercase, abstract, and concrete words. Responses were made by button-press with the right hand. Immediately following the scanning session, women performed a verbal memory task, in which they were presented a list of words, half of which were previously viewed during the verbal processing task during the scanning session. Women were required to indicate whether or not they recalled seeing each of the words during the previous task. To minimize practice effects between placebo and treatment scans, two distinct sets of words were used and counterbalanced to order presented.

Visual working memory was evaluated during the same scanning session using a validated Visual Delayed-Matching-to-Sample task (Lencz et al., 2003; Phillips, 1974). The visual stimuli consisted of 9x9 grids, with 40 squares darkened into a random pattern, presented under three conditions: 1) matching to sample, 2) 1-second delayed-matching-to-sample, or 3) 4-second delayed-matching-to-sample. In the matching condition, a target stimulus was
presented simultaneously with two test items. Matches were indicated by button-press. Stimuli were presented for 3 seconds, followed by a 7-second fixation cross between items. The delay conditions assessed visual working memory: the target stimulus was presented alone for 1.5 seconds, followed by a 1-second or 4-second delay. After the delay, test items were presented for 3 seconds, during which time women indicated the match. Four blocked trials from each condition were counterbalanced over three 6-minute runs, for 180 total scans with a 2-second interscan interval. E-Prime software (Psychology Software Tools Inc., Pittsburgh) controlled the stimulus presentation timing.

Visual and verbal tasks were presented on a computer monitor through radio frequency-shielded goggles mounted to a headcoil (Resonance Technology Inc., Northridge, CA). To minimize performance differences, participants practiced tasks prior to the scanning session until they achieved at least 70% accuracy.

2.7 fMRI acquisition and processing

Scans were acquired using a 3T whole-body MRI scanner (General Electric, Milwaukee) equipped with a standard head coil. Anatomical images were acquired axially with a spoiled gradient recalled (SPGR) three-dimensional volumetric acquisition [repetition time (TR)=9.6, echo time (TE)=3.3, inversion recovery preparation=200 millisecond, flip angle (FA)=17°, bandwidth=15.63, 24cm field of view (FOV), 1.5mm slice thickness, 106–110 slices, 256 x 256 matrix, 2 excitations]. fMRI acquisition was sensitized for the blood oxygen level-dependent (BOLD) effect using a T2* weighted single-shot spiral pulse sequence with 32 oblique-axial slices prescribed to be parallel to the anterior commissure – posterior commissure line (spiral gradient echo, TE=25, TR=2000, FA=60°, 4mm-thick contiguous slices, 24cm FOV, 64x64 image matrix). Image reconstruction included steps to remove distortions caused by magnetic field inhomogeneity and other sources of misalignment to the structural data (Noll et al., 1999). Slices were sinc-interpolated in time to correct for the staggered slice-acquisition sequence (Acquirre and D’Esposito, 1999). The first four volumes of each run were discarded to remove magnetic saturation effects, remaining images were realigned to eliminate movement artifacts, and realignment parameters were examined to ensure head movements did not exceed 2mm (Friston et al., 1995). Anatomical and functional images were coregistered to each other through rigid body affine transformation using a mutual information algorithm (Meyer et al., 1997). Each T1-weighted MRI was spatially normalized via linear and nonlinear warping to a standardized template (International Consortium for Brain Mapping/Montreal Neurological Institute (ICBM/MNI), Quebec). The transformation matrix was applied to functional images, and a three-dimensional Gaussian smoothing kernel set at 8mm full width half maximum was applied to accommodate residual anatomical variability and improve signal-to-noise ratios.

2.8 fMRI data analysis

fMRI data analyses were conducted using the general linear model (GLM) in SPM8 (statistical parametric mapping) imaging software (Wellcome Department of Cognitive Neurology, London). For the first-level analysis, contrast images were generated for each subject to assess activation differences between task conditions. For verbal processing task image analysis, initial contrast images (deep encoding–shallow encoding) were subtracted to
isolate the deep encoding component. For visual memory task image analysis, contrast images were similarly generated to isolate the visual working memory component (4-second delay – 1-second delay). We assessed effects of treatment using these final deep verbal encoding and visual working memory components. We performed initial 1 sample $t$ tests to evaluate the effects of the tasks in our study population, including regions as significant with a false discovery rate (FDR)-corrected $p>0.05$. To compare drug effects, we extracted beta values from these regions and calculated percent signal change for subsequent analyses in SPSS (IBM, Armonk, NY). To fully assess the prefrontal and hippocampal components of working memory circuitry, in addition to those regions meeting significance criteria in the 1 sample $t$ tests, we extracted beta values bilaterally from the hippocampus for both tasks, and from the prefrontal cortex from the visual working memory task, based on peak activation during the task. Paired $t$ tests were performed using extracted data to compare regional activation patterns during the tasks after placebo and estrogen or progesterone treatment. Using data from preliminary studies, we performed sample/size power calculations for alternatives to the null hypothesis (no main effects of treatment). Sample sizes of $N=12$ in each group were considered sufficient to ensure powers $>95\%$ for these analyses, given that the mean effect of task on the BOLD response is expected to be 2%, within subject variability, expressed by the coefficient of variation being conservatively estimated at CV= 0.05, between-subject variability, at CV= 0.10, and the differences between treatment groups between 1.2 and 1.6%.

**Results**

**3.1 Demographics**

Demographic and clinical characteristics are provided in Table 2, and were similar between women randomized to either estrogen or progesterone treatment groups.

**3.2 Neuropsychological assessment of verbal and visual cognition**

Neuropsychological measures of cognitive function indicate similar scores after estrogen or progesterone treatment compared to after placebo for both verbal and visual domains. While much larger sample sizes are typically required to detect subtle differences in cognitive function (Cohen, 1988), we found a statistically significant improvement in verbal working memory after progesterone treatment (Table 3).

**3.3 Neuroimaging task performance**

Performance on the visual processing task did not differ between estrogen and progesterone treatment groups, and was similar after both treatment and placebo arms (Table 4). For the estrogen group, mean accuracy for the shallow processing condition was $99 \pm 2$ (SD)$\%$ after estrogen treatment, and $98 \pm 2\%$ after the placebo arm ($p=0.255$), and mean accuracy for the deep processing condition was $85 \pm 8\%$ after estrogen treatment, and $89 \pm 5\%$ after placebo ($p=0.058$). For the progesterone group, mean accuracy for the shallow processing condition was $98 \pm 2\%$ after progesterone treatment, and $97 \pm 5\%$ after placebo ($p=0.559$), and mean accuracy for the deep processing condition was $87 \pm 7\%$ after progesterone, and $84 \pm 11\%$ after placebo ($p=0.368$). For both groups, memory of words shown during the verbal processing task was slightly better after active treatment compared to placebo, but
differences did not reach statistical significance. Out of 192 words presented, estrogen treated women correctly identified 49 ± 22% of previously viewed words, compared to 37 ± 18% of words after placebo (94 and 52 words, respectively; p=0.092), and progesterone treated women correctly identified 45 ± 20% of words, compared to 33 ± 16% after placebo (86 and 63 words; p=0.270).

Performance on the visual working memory task was also similar between estrogen and progesterone groups. For the estrogen group, accuracy on the match condition was 83 ± 11% after estrogen, and 81 ± 8% after placebo (p=0.687), accuracy on the 1 second delay condition was 91 ± 4% after estrogen and 91 ± 7% after placebo (p=0.871), and accuracy on the 4 second delay condition was 88 ± 8% after estrogen and 89 ± 7% after placebo (p=0.771). For the progesterone group, match condition accuracy was 86 ± 8% after progesterone and 84 ± 8% after placebo (p=0.546), accuracy on the 1 second delay condition was 90 ± 6% after progesterone and 90 ± 8% after placebo (p=0.398), and accuracy on the 4 second delay condition was 90 ± 8% after progesterone and 88 ± 9% after placebo (p=0.480).

3.4 Regional activity during deep verbal encoding and visual working memory

**Verbal task**—We initially determined whole-sample task effects (1 sample T test) for both cognitive tasks (Table 5), using images collected after placebo treatment. For the verbal processing task, we found fMRI BOLD effects in the inferior frontal cortex (MNI coordinates $x,y,z$ (mm) 34, 20, −8; $Z=4.85; P=0.000, P_{FDR}=0.008$ right; −34, 22, −8; $Z=5.22, P=0.000, P_{FDR}=0.002$ left), prefrontal cortex (42, 12, 25; $Z=4.70; P=0.000, P_{FDR}=0.011$ right, −44, 12, 26, $Z=5.68; P=0.000, P_{FDR}=0.001$ left), and superior frontal cortex (−4, 22, 48; $Z=5.91; P=0.000, P_{FDR}<0.001$).

**Visual task**—Task effects for the visual working memory task were found in the posterior cingulate (12, −48, 20; $Z=5.93; P=0.000, P_{FDR}<0.005$).

3.5 Effects of hormone treatment on regional activity during cognitive tasks

**Verbal processing task**—We performed additional analyses on extracted beta values to assess the effects of active estrogen or progesterone treatment on regional activation patterns compared to placebo on regions found significant in the whole-brain 1 sample $t$ test (paired T test; Table 6). For the verbal processing task, we found that estrogen treatment was associated with greater regional activation in the left prefrontal cortex compared to placebo (−44, 48, 2; $P=0.006$), and decreased activation in the left hippocampus (−26, −34, −4; $P=0.037$). Progesterone treatment was associated with decreased activation in the right prefrontal cortex (42, 12, 24; $P=0.014$).

**Visual working memory task**—For the visual working memory task, activation did not differ between estrogen and placebo treatments in any regions (paired $T$ test; Table 6). Progesterone treatment was associated with greater activation in the left prefrontal cortex (−38, 32, 22; $P=0.001$) and the right hippocampus (34, −6, −26; $P=0.003$), compared to placebo. Placebo treatment was not associated with greater activation than progesterone in any region during the visual working memory task.
4. Discussion

The effects of postmenopausal hormone treatment on cognitive outcomes are inconsistent in the literature. Emerging evidence suggests that cognitive effects are influenced by specific hormone formulations, and that progesterone is more likely to be associated with positive outcomes than its synthetic counterparts (L’Hermite, 2013; Simon, 2012). There are very few studies of unopposed progesterone in postmenopausal women, and none that use functional neuroimaging, a sensitive measure that can detect neurobiological changes that precede measurable differences in behavior (Miller et al., 2008; Woodard et al., 2010). In the current study, both estrogen and progesterone treatments were associated with changes in activation patterns during verbal processing compared to placebo, in regions associated with verbal processing and encoding. Conversely, only progesterone was associated with changes in regional brain activation patterns during a visual memory task. Both treatments were associated with a sizeable (but not statistically significant in this relatively small sample) increase in number of words remembered following the verbal task performed during the fMRI scanning session, while only progesterone was associated with improved neuropsychological measures of verbal working memory compared to placebo. In contrast to previous studies that have found negative cognitive effects of treatment with synthetic progestins, our results point to potential cognitive benefits of both estrogen and progesterone.

4.1 Estrogen

In contrast to the variable outcomes associated with combined hormone treatments, investigations of unopposed estradiol have largely found positive or neutral effects on verbal cognition, although results from studies using conjugated equine estrogens (CEE) have been less conclusive (Dumas et al., 2010; Espeland et al., 2013; Gleason et al., 2006; Gorenstein et al., 2011; Maki, 2012; Maki and Sundermann, 2009; Sherwin, 2012; Sherwin and Grigorova, 2011; Wroolie et al., 2011). We expected that short-term estradiol administration would positively impact verbal cognition. While we did not detect differences in verbal ability between the estrogen and placebo treatment arms, we saw increased activation after estrogen treatment in the left prefrontal cortex during the verbal processing task, a region associated with verbal processing. This may reflect more efficient encoding, as women remembered more words from the verbal task after estrogen than placebo, although this difference did not reach statistical significance. Previous studies of postmenopausal estrogen use have found correlations between increased prefrontal metabolism and verbal cognitive performance (Silverman et al., 2011). We did not find any effect of estrogen on visual memory. Our results are aligned with previous evidence that menopausal estrogen treatment increases prefrontal activation, and may selectively benefit prefrontal cognitive processes (Joffe et al., 2006). Several mechanisms are likely involved, including modulation of interactions with neurotransmitter systems, growth factors, and dendritic spine density (Shanmugan and Epperson, 2014). Studies in rats have found that estrogenic effects may be mediated through extra-nuclear estrogen regulation of excitatory glutamatergic synapse formation, and estrogen may also regulate brain derived neurotrophic factor (BDNF) levels via nuclear receptor activity (Khan et al., 2013; Luine and Frankfurter, 2013). However there is evidence both for and against estrogen-mediated increases in prefrontal dendritic spine...
density in non-human primate models of menopause, and it is likely that estrogen mediates neuronal activity in this region through a combination of nuclear and extra-nuclear actions promoting interactions with several neurotransmitter systems and growth factors (Hao et al., 2006; Ohm et al., 2012; Shanmugan and Epperson, 2014).

### 4.2 Progesterone

Synthetic progestins have been consistently associated with reductions in verbal ability, regardless of treatment timing or postmenopausal age group (Coker et al., 2010; Maki et al., 2007; Resnick et al., 2004; Resnick et al., 2006). Verbal impact appears to be heavily influenced by progestin formulation, however, and progestins with less anti-estrogenic effects have been shown to have neutral or positive impact on verbal benefits conferred by estradiol treatment (Spark and Willis, 2012). Progesterone itself has neurobiological actions unique from those of progestins, including neurotrophic and neuroprotective effects in animal models, however there is little known about the potential effects of progesterone on postmenopausal cognition (Pluchino et al., 2009; Schussler et al., 2008). Progesterone has been shown to modulate cognitive function in aging rats through actions in the hippocampus and prefrontal cortex (Camacho-Arroyo et al., 2011; Frye and Walf, 2008; Paris et al., 2011). While there is ample evidence for estrogenic regulation of prefrontal dendritic spine density, progesterone mechanisms in this region are less clear. A study in ovariectomized rats found that 18 weeks of progesterone, but not estradiol, increased expression of glial fibrillary acidic protein and microtubule-associated proteins, which are associated with neurotrophic actions including neuronal transport and plasticity as well as dendrite stability and extension (Camacho-Arroyo et al., 2011).

We found that verbal working memory was improved after progesterone treatment compared to placebo, and we found increased activation in the left prefrontal cortex and right hippocampus during the visual working memory task. Interestingly, lateralization of hippocampal activation reflects previous evidence of left hippocampal regulation of verbal encoding and right hippocampal regulation of visual encoding (Papanicolaou et al., 2002). The hippocampus may be particularly sensitive to neuromodulatory effects of progesterone. In rats, progesterone has been shown to protect hippocampal neurons from traumatic injury in culture and live animals, with a corresponding improvement in cognitive outcomes following injury (Si et al., 2013; Si et al., 2014). This sensitivity to progesterone has been attributed to antioxidant effects or increased hippocampal neurogenesis in rats (Chan et al., 2014; He et al., 2011). Interestingly, a recent study found that progesterone or progesterone with estradiol increased hippocampal neurogenesis and neuronal survival in female rats, while MPA or estradiol alone reduced neurogenesis and neuronal survival, which could explain the increases in activation following progesterone but not estrogen treatment observed in the present study (Chan et al., 2014). Similarly, a study of endogenous hormones found that serum progesterone, but not estrogen concentration was associated with verbal cognitive function (Henderson et al., 2013). Previous studies in menopausal women have determined that cognitive benefits of hormone use may be attributed to estrogenic effects on the hippocampus, but these effects may occur with longer term use than in the current study (Maki et al., 2011; Resnick et al., 2009), and may also be mediated...
by estrogen-inducible progesterone receptors (McEwen et al., 2012; Woolley and McEwen, 1993).

4.3 Limitations & strengths

Conclusions drawn from this pilot study are limited by the small sample size, however the crossover design allowed maximal statistical power given the number of subjects. Two of the women included in the study had hysterectomy and bilateral oophorectomy. The abrupt nature of surgical menopause may have distinct neurobiological effects from natural menopause, however these women met all other screening criteria, and were equally distributed between the treatment groups, minimizing any potential impacts of such differences. An additional limitation is the relatively short treatment duration. Future studies with a larger sample size and longer treatment duration would likely result in a more accurate assessment of changes in verbal and visual cognitive function, in addition to the changes we detected using more sensitive neuroimaging measures of the underlying neurobiology, and would also allow for comparisons between estrogen and progesterone treatments. While cognitive capacity is most directly measured through neuropsychological assessments, these behavioral measures tend to be variable and require a large sample size to detect subtle differences in similar populations. Functional MRI can detect early structural or processing pathway alterations that have been shown to precede measurable differences in cognitive function, providing a measure of early-stage neurobiological differences (Miller et al., 2008; Woodard et al., 2010). Increasing the length of treatment in future studies may allow a more complete assessment of differences in cognitive ability between treatment groups, and detection of relationships between task performance and regional brain activation patterns. Strengths include randomized design, and separation of estrogen and progesterone treatments, allowing delineation of the neurobiological effects of each hormone. All women were within 38 months of their final menstrual period, a time point falling within the ‘critical window’ when maximal benefit of postmenopausal hormone treatment is thought to occur (Maki, 2013; Whitmer et al., 2011).

4.4 Conclusions

Despite uncertainty about the potential cognitive benefits of postmenopausal hormone use, as women continue to live longer and healthier lives beyond the end of their reproductive years, hormones will continue to be prescribed for symptomatic indications. Understanding the differential effects of estrogen and progesterones a critical step leading to more effective therapies that are most likely to impart a cognitive benefit. Our results point to potential cognitive impacts of both estrogen and progesterone.

Acknowledgements

This work was supported by NIH grants R21AG031951 and 5K01MH095920, the NIH CTSA grant 2UL1TR000433-06 for MCRU services, and by the Phil F. Jenkins Foundation (JKZ). We thank the University of Michigan fMRI laboratory, Anne Tkaczyk for study coordination, and especially the women who participated in our study.

Funding sources had no involvement in study design, data collection, analysis, or interpretation, in writing the report, or in the decision to submit the article for publication.
References


Acquirre, C.; D’Esposito, M. Experimental design for brain fMRI. Heidelberg: Springer-Verlag; 1999.


• Recently menopausal women were given estradiol or progesterone for 90 days
• Both groups had more regional activity during verbal processing compared to placebo
• For visual memory, women on progesterone had more regional activity versus placebo
• Only progesterone was associated with better verbal working memory versus placebo
• Progesterone treatment may impart a cognitive benefit
Figure 1. Regional activation during verbal processing (top) and verbal working memory (bottom) tasks in all women after placebo treatment arm

Whole-brain 1 sample t tests revealed bilateral inferior frontal, prefrontal, and superior frontal cortical activation during a verbal processing task, and posterior cingulate activation during a visual working memory task. Scale bar indicates T score.
<table>
<thead>
<tr>
<th>Enroll  (29)</th>
<th>Estrogen (16)</th>
<th>Progesterone (13)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Estrogen</td>
<td>Progesterone</td>
</tr>
<tr>
<td></td>
<td>Placebo</td>
<td>Placebo</td>
</tr>
<tr>
<td></td>
<td>12 weeks</td>
<td>10 days</td>
</tr>
<tr>
<td></td>
<td>Progesterone</td>
<td>Placebo</td>
</tr>
<tr>
<td></td>
<td>Placebo</td>
<td>Placebo</td>
</tr>
<tr>
<td></td>
<td>↑</td>
<td></td>
</tr>
<tr>
<td></td>
<td><em>fMRI &amp; behavioral testing</em></td>
<td><em>fMRI &amp; behavioral testing</em></td>
</tr>
</tbody>
</table>
## Table 2

Demographic & baseline assessments

<table>
<thead>
<tr>
<th></th>
<th>Whole Group (29)</th>
<th>Estrogen (16)</th>
<th>Progesterone (13)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean ± SD</td>
<td>Mean ± SD</td>
<td>Mean ± SD</td>
</tr>
<tr>
<td>Age (yrs)</td>
<td>51.52 ± 2.52</td>
<td>51.81 ± 2.23</td>
<td>51.15 ± 2.88</td>
</tr>
<tr>
<td>Education (yrs)</td>
<td>16.40 ± 2.80</td>
<td>15.75 ± 2.82</td>
<td>17.19 ± 2.67</td>
</tr>
<tr>
<td>Months since LMP</td>
<td>18.41 ± 9.63</td>
<td>19.81 ±9.57</td>
<td>16.69 ± 9.80</td>
</tr>
<tr>
<td>Estradiol (pg/mL)</td>
<td>16.24 ± 6.40</td>
<td>17.75 ± 6.30</td>
<td>14.38 ± 6.25</td>
</tr>
<tr>
<td>FSH (mIU/mL)</td>
<td>87.19 ± 29.11</td>
<td>92.34 ±33.85</td>
<td>80.86 ± 21.60</td>
</tr>
<tr>
<td>BMI (kg/m(^2))</td>
<td>27.69 ± 5.17</td>
<td>27.13 ± 6.13</td>
<td>28.38 ± 3.80</td>
</tr>
<tr>
<td>Measure</td>
<td>Estrogen Mean ± SD</td>
<td>Placebo Mean ± SD</td>
<td>p</td>
</tr>
<tr>
<td>---------------------------------</td>
<td>-------------------</td>
<td>-------------------</td>
<td>-----</td>
</tr>
<tr>
<td>Verbal Learning (HVLT-R Total)</td>
<td>29.7 ± 3.2</td>
<td>28.5 ± 4.2</td>
<td>0.10</td>
</tr>
<tr>
<td>Verbal Recall (HVLT-R Delayed Recall)</td>
<td>10.9 ± 1.2</td>
<td>10.5 ± 2.0</td>
<td>0.19</td>
</tr>
<tr>
<td>Verbal Working Memory (composite score)</td>
<td>96 ± 8</td>
<td>99 ± 10</td>
<td>0.28</td>
</tr>
<tr>
<td>Visual Learning (BVMT-R Total)</td>
<td>23.7 ± 5.5</td>
<td>22.9 ± 7.8</td>
<td>0.56</td>
</tr>
<tr>
<td>Visual Recall (BVMT-R Delayed Recall)</td>
<td>9.3 ± 2.2</td>
<td>9.3 ± 2.3</td>
<td>0.88</td>
</tr>
<tr>
<td>Visual Working Memory (WMS-III Spatial Span)</td>
<td>18.1 ± 1.9</td>
<td>18.3 ± 2.5</td>
<td>0.77</td>
</tr>
</tbody>
</table>

The Verbal Working Memory Composite represents the averaged sample-based standard scores (M=100, SD=15) of WAIS-III Digit Span, Brown Peterson, and WAIS-III Letter Number Sequencing. All other values represent raw scores.
Table 4

<table>
<thead>
<tr>
<th></th>
<th>Estrogen</th>
<th>Placebo</th>
<th>Progesterone</th>
<th>Placebo</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean ± SD</td>
<td>Mean ± SD</td>
<td>Mean ± SD</td>
<td>Mean ± SD</td>
<td></td>
</tr>
<tr>
<td><strong>Verbal Processing</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Shallow Processing</td>
<td>Accuracy</td>
<td>0.99 ± 0.02</td>
<td>0.98 ± 0.02</td>
<td>0.255</td>
<td>0.97 ± 0.05</td>
</tr>
<tr>
<td></td>
<td>Reaction Time</td>
<td>0.79 ± 0.08</td>
<td>0.76 ± 0.06</td>
<td>0.084</td>
<td>0.77 ± 0.07</td>
</tr>
<tr>
<td>Deep Processing</td>
<td>Accuracy</td>
<td>0.85 ± 0.08</td>
<td>0.89 ± 0.05</td>
<td>0.058</td>
<td>0.87 ± 0.07</td>
</tr>
<tr>
<td></td>
<td>Reaction Time</td>
<td>1.08 ± 0.11</td>
<td>1.04 ± 0.11</td>
<td>0.213</td>
<td>1.07 ± 0.10</td>
</tr>
<tr>
<td><strong>Verbal Recognition</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Shallow Processing</td>
<td>Accuracy</td>
<td>0.58 ± 0.30</td>
<td>0.42 ± 0.26</td>
<td>0.228</td>
<td>0.54 ± 0.28</td>
</tr>
<tr>
<td></td>
<td>Reaction Time</td>
<td>0.93 ± 0.09</td>
<td>0.94 ± 0.10</td>
<td>0.609</td>
<td>0.93 ± 0.11</td>
</tr>
<tr>
<td>Deep Processing</td>
<td>Accuracy</td>
<td>0.41 ± 0.23</td>
<td>0.31 ± 0.15</td>
<td>0.087</td>
<td>0.36 ± 0.18</td>
</tr>
<tr>
<td></td>
<td>Reaction Time</td>
<td>0.94 ± 0.11</td>
<td>0.96 ± 0.08</td>
<td>0.581</td>
<td>0.93 ± 0.09</td>
</tr>
<tr>
<td><strong>Visual Working Memory</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Match</td>
<td>Accuracy</td>
<td>0.83 ± 0.11</td>
<td>0.81 ± 0.08</td>
<td>0.687</td>
<td>0.86 ± 0.08</td>
</tr>
<tr>
<td></td>
<td>Reaction Time</td>
<td>1.81 ± 0.23</td>
<td>1.79 ± 0.14</td>
<td>0.815</td>
<td>1.79 ± 0.12</td>
</tr>
<tr>
<td>1s Delay</td>
<td>Accuracy</td>
<td>0.91 ± 0.04</td>
<td>0.91 ± 0.07</td>
<td>0.871</td>
<td>0.87 ± 0.06</td>
</tr>
<tr>
<td></td>
<td>Reaction Time</td>
<td>1.53 ± 0.13</td>
<td>1.52 ± 0.13</td>
<td>0.950</td>
<td>1.41 ± 0.16</td>
</tr>
<tr>
<td>4s Delay</td>
<td>Accuracy</td>
<td>0.88 ± 0.08</td>
<td>0.89 ± 0.07</td>
<td>0.771</td>
<td>0.90 ± 0.08</td>
</tr>
<tr>
<td></td>
<td>Reaction Time</td>
<td>1.60 ± 0.15</td>
<td>1.58 ± 0.09</td>
<td>0.643</td>
<td>1.49 ± 0.17</td>
</tr>
</tbody>
</table>
Table 5

Regions activated during verbal processing and visual working memory tasks

<table>
<thead>
<tr>
<th>Verbal Processing Effect of Task: all subjects (placebo arm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Region</td>
</tr>
<tr>
<td>--------</td>
</tr>
<tr>
<td>R inferior frontal cortex</td>
</tr>
<tr>
<td>L inferior frontal cortex</td>
</tr>
<tr>
<td>R prefrontal cortex</td>
</tr>
<tr>
<td>L prefrontal cortex</td>
</tr>
<tr>
<td>B superior frontal cortex</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Visual Working Memory Effect of Task: all subjects (placebo arm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Region</td>
</tr>
<tr>
<td>--------</td>
</tr>
<tr>
<td>Posterior cingulate</td>
</tr>
</tbody>
</table>
## Table 6

Treatment effects on regional activity during verbal processing and visual working memory tasks

<table>
<thead>
<tr>
<th>Region</th>
<th>Estrogen Mean ± SD</th>
<th>Placebo Mean ± SD</th>
<th>p</th>
<th>Progesterone Mean ± SD</th>
<th>Placebo Mean ± SD</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>R inferior frontal cortex</td>
<td>0.33 ± 0.18</td>
<td>0.28 ± 0.18</td>
<td>0.565</td>
<td>0.12 ± 0.15</td>
<td>0.15 ± 0.16</td>
<td>0.714</td>
</tr>
<tr>
<td>L Inferior frontal cortex</td>
<td>0.21 ± 0.18</td>
<td>0.27 ± 0.14</td>
<td>0.461</td>
<td>0.08 ± 0.25</td>
<td>0.15 ± 0.15</td>
<td>0.343</td>
</tr>
<tr>
<td>R prefrontal cortex</td>
<td>0.28 ± 0.23</td>
<td>0.28 ± 0.21</td>
<td>0.984</td>
<td>0.00 ± 0.12</td>
<td>0.19 ± 0.19</td>
<td>0.014</td>
</tr>
<tr>
<td>B superior frontal cortex</td>
<td>0.31 ± 0.18</td>
<td>0.14 ± 0.25</td>
<td>0.006</td>
<td>0.04 ± 0.29</td>
<td>0.13 ± 0.30</td>
<td>0.345</td>
</tr>
<tr>
<td>L hippocampus</td>
<td>0.36 ± 0.15</td>
<td>0.36 ± 0.21</td>
<td>0.919</td>
<td>0.19 ± 0.09</td>
<td>0.27 ± 0.16</td>
<td>0.189</td>
</tr>
<tr>
<td>R hippocampus</td>
<td>−0.08 ± 0.12</td>
<td>0.00 ± 0.14</td>
<td>0.578</td>
<td>−0.07 ± 0.13</td>
<td>0.01 ± 0.07</td>
<td>0.091</td>
</tr>
<tr>
<td>L hippocampus</td>
<td>−0.08 ± 0.10</td>
<td>0.00 ± 0.13</td>
<td>0.037</td>
<td>−0.08 ± 0.13</td>
<td>−0.04 ± 0.07</td>
<td>0.437</td>
</tr>
<tr>
<td>R prefrontal cortex</td>
<td>0.31 ± 0.29</td>
<td>0.23 ± 0.30</td>
<td>0.549</td>
<td>0.37 ± 0.31</td>
<td>0.29 ± 0.25</td>
<td>0.493</td>
</tr>
<tr>
<td>L prefrontal cortex</td>
<td>0.18 ± 0.26</td>
<td>0.16 ± 0.35</td>
<td>0.853</td>
<td>0.42 ± 0.22</td>
<td>−0.02 ± 0.31</td>
<td>0.001</td>
</tr>
<tr>
<td>R hippocampus</td>
<td>0.33 ± 0.23</td>
<td>0.24 ± 0.44</td>
<td>0.435</td>
<td>0.49 ± 0.33</td>
<td>0.16 ± 0.26</td>
<td>0.003</td>
</tr>
<tr>
<td>L hippocampus</td>
<td>0.25 ± 0.27</td>
<td>0.25 ± 0.34</td>
<td>0.993</td>
<td>0.36 ± 0.20</td>
<td>0.38 ± 0.34</td>
<td>0.806</td>
</tr>
<tr>
<td>Posterior cingulate</td>
<td>0.42 ± 0.35</td>
<td>0.30 ± 0.14</td>
<td>0.209</td>
<td>0.45 ± 0.33</td>
<td>0.31 ± 0.17</td>
<td>0.201</td>
</tr>
</tbody>
</table>

Regional fMRI BOLD signal change patterns were compared between estrogen or progesterone and placebo treatment arms during verbal encoding and visual working memory tasks, using beta values extracted from regions of interest. Both hormone treatments were associated with greater regional activation than placebo treatment in the indicated regions during the verbal processing task, while only progesterone was associated with greater regional activity than placebo during the visual working memory task.