Sex differences in the effect of acute tryptophan depletion on declarative episodic memory: A pooled analysis of nine studies

Anke Sambetha,b,* Arjan Bloklanda,b, Catherine J. Harmerc, Tessa O.C. Kilkensb, Pradeep J. Nathand, Richard J. Portere, Jeroen A.J. Schmittf, Bart Scholtissenb, Sjacco Sobczakb, Allan H. Youngg, Wim J. Riedela,h,i

a Department of Neurocognition, Faculty of Psychology, Maastricht University, P.O. Box 616, 6200 MD Maastricht, The Netherlands
b Brain and Behaviour Institute, Maastricht University, The Netherlands
c Department of Experimental Psychology, University of Oxford, South Parks Road, Oxford OX1 3UD, UK
d Department of Physiology and Monash Centre for Brain and Behaviour, Monash University, Clayton 3800, Vic., Australia
e Nestlé Research Centre, Nutrition and Health Department, P.O. Box 33, Vers-chez-les-Blanc, 1000 Lausanne 26, Switzerland
f Department of Psychiatry, University of British Columbia, 2255 Wesbrook Mall, Vancouver, BC, Canada V6T 2A1
g GlaxoSmithKline R&D, Clinical Pharmacology & Discovery Medicine—Psychiatry, Addenbrookes Hospital, Cambridge, UK
h Nestlé Research Centre, Nutrition and Health Department, P.O. Box 33, Vers-chez-les-Blanc, 1000 Lausanne 26, Switzerland

Received 21 August 2006; received in revised form 27 November 2006; accepted 27 November 2006

Abstract

Acute tryptophan depletion (ATD) studies have shown that serotonin plays a role in learning and memory processes. In this study, we performed a pooled analysis of nine ATD studies in order to examine the nature of the memory-impairing effects of ATD and mediating factors, such as gender, age and vulnerability for disease in which disturbed serotonin was hypothesized to play a role. All studies that were used in this pooled analysis assessed declarative episodic memory using a verbal learning task paradigm. Immediate recall, delayed recall, and delayed recognition scores were examined. A total of 211 participants were included in the analysis. The analysis revealed that ATD impaired not only delayed recall, but also immediate recall. The ATD-induced impairments were larger in females than in males. Furthermore, ATD did not interact with any other serotonergic vulnerability and age. This suggests that the only factor that actually has the properties of a serotonergic vulnerability factor for declarative memory performance is female gender. The findings provide further support for a critical role of serotonin in declarative episodic memory.

Keywords: Acute tryptophan depletion; Age; Declarative memory; Episodic memory; Serotonergic vulnerability; Sex differences

Contents

1. Introduction ................................................................. 517
2. Materials and methods .................................................. 518
   2.1. Participants .......................................................... 518
   2.2. Study designs ....................................................... 519
   2.3. Treatment .......................................................... 519
   2.4. Assessments ......................................................... 519

*Corresponding author. Department of Neurocognition, Faculty of Psychology, Maastricht University, P.O. Box 616, NL-6200 MD Maastricht, The Netherlands. Tel.: + 31 433881757; fax: + 31 433884125.
E-mail address: anke.sambeth@psychology.unimaas.nl (A. Sambeth).

© 2006 Elsevier Ltd. All rights reserved.

doi:10.1016/j.neubiorev.2006.11.009
1. Introduction

Serotonergic neurotransmission is involved in the regulation of mood, but also plays a role in modulating cognitive processes such as learning and memory (Meneses, 1999; Buhot et al., 2000; Schmitt et al., 2006). One method to examine the function of serotonin neurotransmission in the brain in the context of learning and memory is acute tryptophan depletion (ATD). ATD causes a temporary global reduction of serotonin (5-HydroxyTryptophan; 5-HT) synthesis in the brain by decreasing the availability of its precursor L-tryptophan (TRP) (Moore et al., 2000; Sobczak et al., 2002; Harrison et al., 2004; Kilkens et al., 2004; Scholtissen et al., 2006).

Depressed patients show deficits on neuropsychological tests especially of learning, memory and executive function (McAllister-Williams et al., 1998; Austin et al., 2001; Porter et al., 2003a). It has been argued that ATD-induced cognitive impairments offer an objective means of studying this aspect of the pathophysiology of depression (McAllister-Williams et al., 2002; Riedel et al., 2002).

Cognitive effects of ATD, in particular an impairment of memory for recently acquired information, have been observed in healthy subjects with and without a vulnerability for mood disorders (Park et al., 1994; Riedel et al., 1999; Schmitt et al., 2000; Shansis et al., 2000; Murphy et al., 2002; Riedel et al., 2002; Rubinsztein et al., 2001; Sobczak et al., 2002). In this paper, we will focus on several aspects of the nature of the memory-impairing effects of ATD and mediating factors that have hitherto been unresolved, such as gender, age and vulnerability for affective disorders. In particular, the focus is on the role of 5-HT in declarative memory processes that are thought to depend on structures in the medial temporal lobe and the midline diencephalon (Squire, 2004). Storing information into memory involves three stages at large: encoding, consolidation, and retrieval. Serotonin may specifically and differentially influence declarative memory processes, especially the encoding and consolidation phases (Meneses, 1999; Meneses and Hong, 1999; Levkovitz et al., 2002; Meeter et al., 2006; Reynolds et al., 2006; Buhot et al., 2000).

Learning and memory tasks that were studied extensively in humans after ATD, probed declarative episodic memory using a verbal learning task paradigm. It has been shown that delayed recall and recognition of words was impaired when learning occurred after ATD (Riedel et al., 1999; Schmitt et al., 2000; McAllister-Williams et al., 2002; Harrison et al., 2004; Kilkens et al., 2004; Scholtissen et al., 2006; Sobczak et al., 2002), while retrieval from semantic memory was not impaired after ATD or even improved (Schmitt et al., 2000; McAllister-Williams et al., 2002), suggesting that specifically consolidation of episodic memory is impaired after ATD (Riedel et al., 1999; Schmitt et al., 2000; Sobczak et al., 2002; Harrison et al., 2004; Kilkens et al., 2004; Scholtissen et al., 2006).
Effects on immediate recall, which may primarily reflect effects on the encoding process, have not been found as extensively as delayed recall effects. Kilkens et al. (2004) and Porter and colleagues (2005) showed modest but statistically reliable reductions in immediate recall after ATD, whereas Scholtissen et al. (2006) only found a marginal effect. Hayward and co-workers (2005) showed that immediate recall was impaired by ATD in recovered depressed subjects, but not in healthy controls. Most of the other studies using verbal learning reported small and non-significant reductions in immediate recall after ATD as well (Riedel et al., 1999; Schmitt et al., 2000; Sobczak et al., 2002; Hughes et al., 2003; Kilkens et al., 2004). One aspect of the immediate recall score that has not been given attention so far after ATD is the steepness of the learning curve, that is, the increment in the number of remembered words after multiple presentations of a word list. This measure can, however, be used as another indicator for possibly disturbed working memory processes after ATD. In the present study, it is examined whether ATD affects the steepness of learning.

ATD decreases the rate of serotonin synthesis in the brain, which can be studied using PET tracers (Nishikawa et al., 2005). This effect is larger in females than in males, as was shown in a PET study using a tracer of tryptophan (Nishizawa et al., 1997). The mood-lowering effects of ATD have been shown to be larger in females as well (Ellenbogen et al., 1996; Smith et al., 1997a; Booij et al., 2002). Consequently, one might expect that ATD effects on memory would also be larger in females than in males. However, ATD studies, in which the influence of gender differences on verbal learning were investigated, failed to show those effects (Riedel et al., 1999; Schmitt et al., 2000). It should be mentioned, however, that none of these studies were statistically adequately powered in terms of sample size, to analyze gender effects.

ATD has been applied several times to investigate whether specified groups of otherwise healthy subjects are more vulnerable to the effects of serotonergic manipulations. Healthy subjects with a positive family history of depression showed significantly more depressed mood after ATD than healthy subjects without such a family history (Klaassen et al., 1999). ATD in depressed patients in remission who had been treated with serotonergic antidepressants induced clinically significant depressive symptoms within 6 h (Delgado et al., 1989, 1991; Lam et al., 1996; Smith et al., 1997b). These symptoms disappeared within 24 h after ATD. Likewise, irritable bowel syndrome (IBS) is a disease of visceral hypersensitivity and is associated with a high co-occurrence of psychiatric symptoms, in particular affective dysregulation (Kilkens et al., 2004). It has been hypothesized that the visceral and affective IBS symptoms are mediated by peripheral as well as central 5-HT dysregulation leading to mood and cognitive symptoms akin to those manifested in depression (Kilkens et al., 2003). Some attempts have been undertaken to investigate if similar vulnerabilities exist with respect to ATD-induced memory impairments. It has been shown that ATD does not lead to greater memory impairment in participants with a serotonergic dysregulation in comparison to healthy controls (Riedel et al., 1999; Sobczak et al., 2002; Kilkens et al., 2004). However, a lack of statistical power may explain these null-results. Therefore, in addition to healthy controls, healthy recovered depressed patients, healthy participants with a first-degree family history of depression and psychiatrically healthy IBS patients were compiled into one group and compared in terms of memory performance to healthy controls. The compiled group is called ‘hypothesized serotonergic vulnerability’ group, since the common element in these groups is their increased probability to develop affective disorders. Despite differences in etiology, they are all hypothesized to have a 5-HT dysregulation. Our hypothesis is that poorer memory performance is linked to a heightened sensitivity of the serotonergic system. That is why memory problems are associated with a number of disorders, in which 5-HT disturbance is thought to play a role.

The justification for pooling these groups into one is the following. According to the functional approach in biological psychiatry, 5-HT disturbances in psychiatric diseases are not linked to one particular disorder. Rather, they are associated with dysregulations of specific psychopathological dimensions (Van Praag et al., 1975, 1987), such as aggression, mood, or memory. In this pooled analysis, the ‘hypothesized serotonergic vulnerability group’ is compiled according to this functional principle to examine dysfunctions in memory processing.

To increase statistical power, a ‘mega-analysis’ or ‘pooled-analysis of original data’ (Olkin, 1995; Booij et al., 2002) can be conducted. In such an analysis, the raw data from several studies are combined within one analysis, rather than measuring effect sizes from several studies, as is done in more commonly known meta-analyses. In this study, we performed a mega-analysis on data of several studies that had investigated the effects of ATD on declarative verbal memory using adapted versions of the Rey Auditory Verbal Learning Test (Rey, 1964). Furthermore, effects of gender, age and a hypothesized serotonergic vulnerability were assessed.

2. Materials and methods

2.1. Participants

After a literature search using Medline and PsycLit sources, various research groups throughout the world were approached. The main criteria were having published one or more studies examining memory effects of ATD using an amino acid mixture as treatment, using a word list learning test as dependent measure of declarative memory, and the use of a within-subject, placebo-controlled design. Characteristics and descriptors of the nine studies that were included in the analysis are presented in Table 1.
The data set for the analysis contained 211 participants in the age range of 18–85 years, of which 137 were healthy controls and 74 were participants with a hypothesized central and/or peripheral serotonergic dysregulation or vulnerability (SDV) (either first-degree family history of unipolar depression, first-degree family history of bipolar depression, recovered depressed patients, or IBS patients. Family history for depression was assessed using the diagnostic criteria of Endicott et al. (1975)).

The gender distribution over all studies was 112 females and 99 males. In order to control for possible phase-dependent variations in mood or cognitive performance in females, several precautions were taken per study. In the study of Hughes et al. (2003), only male participants were included, whereas Sobczak et al. (2002), Harrison et al. (2004), and Kilkens et al. (2004) tested females in the follicular phase of their menstrual cycle. The day of the cycle was registered in the study of Riedel et al. (1999) and, in both the ATD and placebo conditions, only two female participants were tested during the pre-menstrual week. Furthermore, there was no difference between the groups. In the studies of Porter et al. (2000, 2005) and Scholtissen et al. (2006), all participants were post-menopausal.

Common exclusion criteria in the different studies were the use of medication other than oral contraceptives, the use of psychoactive medication, current physical illness, current depression, history of psychiatric illness, and a positive first-degree psychiatric family history. The latter two exclusion criteria were not used with regard to recovered depressed participants and participants with a family history for depression in the studies that examined the hypothesized serotonergic vulnerability. They, however, did apply to the controls in those studies.

### 2.2. Study designs

All studies were using a placebo-controlled, double-blind, cross-over design. The treatment consisted of the administration of an amino acid mixture to induce ATD or a placebo and the order was balanced over test days.

### 2.3. Treatment

In all studies, the tryptophan-free mixture contained 15 amino acids, after the method of Young et al. (1985). See Table 2 for a complete list of the amino acids. The amount of the mixture differed between studies (see Table 1). In the placebo drink, L-tryptophan was added. The amount of L-tryptophan used varied somewhat between studies (see Table 1).

### 2.4. Assessments

#### 2.4.1. The Visual Verbal Learning Test (VVLT)

In all studies, adapted versions of the Rey Auditory Verbal Learning Test (Lezak, 1995) were used. See Table 3.
for a summary of the parameters for each of the studies included in the analysis. In the verbal learning test, either 15 or 30 words (see Table 3) are presented on a computer screen in 1, 3, or 5 trials (see Table 3). Each trial ends with a free recall of the words (immediate recall). If more than three trials were presented, only the first three were used for analysis of immediate recall. Fifteen to thirty minutes after the last trial (see Table 3), the participant is asked to recall as many words as possible without prompting (delayed recall). This is followed by a recognition test, consisting of previously presented words and other but comparable words. In this test, the participant has to respond ‘YES/NO’ or similar as fast as possible to indicate recognition of the word (delayed recognition). The responses to immediate recall, delayed recall, and delayed recognition were used in the analysis.

2.4.2. Blood levels

In the different studies, blood samples were obtained prior to and 4–6.5 h after ingestion of the amino acid mixture in each of the studies. Plasma tryptophan concentrations and/or tryptophan/\(\Sigma\)LNAA (large neutral amino acids) ratios were calculated. The plasma concentrations of tryptophan were significantly decreased after ATD in all studies compared to a baseline measure. TRP decreases varied from 63% to 97% across studies, while decreases in the tryptophan/\(\Sigma\)LNAA ratio varied from 70% to 80%, but this measure was not reported in all studies.

2.4.3. Additional assessments

In most of the studies, additional cognitive measures were taken from the participants. These measures will not be addressed in this analysis. The depression score of the profile of mood states (POMS) and the TRP plasma concentrations, however, will be examined to clarify and interpret the memory results.

2.5. Procedure

Participants arrived at the laboratory after an overnight fast in all of the studies. In most of the studies, but not Riedel et al. (1999), Hughes et al. (2003), Kilkens et al. (2004) a baseline recording was taken first, in which all tasks were performed once. Thereafter, the participants were offered the drink with or without tryptophan, which had to be consumed within at least 2 h. Four to six-and-a-half hours after ingestion, the word learning task was performed and blood samples were taken.\(^1\) At least four days later, the participants returned to the laboratory for the second assessment, in which they received the other drink and performed the assessments.

2.6. Missing data

In some of the studies, one or more variables were missing for one or more participants. In the study of

---

Table 3
Parameters and brief results of the Word Learning Task for each of the studies

<table>
<thead>
<tr>
<th>Reference</th>
<th>Number of words in the list</th>
<th>Number of immediate recall trials</th>
<th>Delay between immediate and delayed recall (min)</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Riedel et al. (1999)</td>
<td>30</td>
<td>3</td>
<td>30</td>
<td>Delayed recall impaired in TRP—condition, recognition impaired after 24 h in TRP—condition</td>
</tr>
<tr>
<td>Schmitt et al. (2000)</td>
<td>30</td>
<td>3</td>
<td>30</td>
<td>Delayed recall and recognition impaired in TRP—condition</td>
</tr>
<tr>
<td>Porter et al. (2000)</td>
<td>15</td>
<td>5</td>
<td>20</td>
<td>No effects of tryptophan depletion on learning and memory processes</td>
</tr>
<tr>
<td>Sobczak et al. (2002)</td>
<td>30</td>
<td>3</td>
<td>30</td>
<td>Delayed recall impaired in TRP—condition</td>
</tr>
<tr>
<td>Hughes et al. (2003)</td>
<td>15</td>
<td>5 (not included in this analysis)</td>
<td>30</td>
<td>No effects of tryptophan depletion on learning and memory processes</td>
</tr>
<tr>
<td>Harrison et al. (2004)</td>
<td>15</td>
<td>1</td>
<td>20</td>
<td>Delayed recall impaired in TRP—condition</td>
</tr>
<tr>
<td>Kilkens et al. (2004)</td>
<td>30</td>
<td>3</td>
<td>30</td>
<td>Both immediate and delayed recall impaired in TRP—condition</td>
</tr>
<tr>
<td>Porter et al. (2005)</td>
<td>15</td>
<td>5</td>
<td>20</td>
<td>Immediate recall impaired in TRP—condition</td>
</tr>
<tr>
<td>Scholtissen et al. (2006)</td>
<td>15</td>
<td>3</td>
<td>25–30</td>
<td>Delayed recall and recognition impaired in TRP—condition</td>
</tr>
</tbody>
</table>

\(^1\) Previous research has shown that maximal reduction of plasma TRP by 75–90% occurs at 4 h and remains low for at least 2 h (Young et al., 1985; Klaassen et al., 1999). Some have argued that approximately a 2 h delay exists between the lowering of TRP and the consequent depletion of brain 5-HT. The metabolite of 5-HT, 5-HIAA, was shown to be lowered maximally between 8–12 h after ingestion of the drink (Carpenter et al., 1998). However, empirical observations have confirmed that ATD-induced memory impairments were no more pronounced at either 9 or 24 h after the drink (Riedel et al., 1999; Schmitt et al., 2000).
Schmitt and colleagues (2000), the delayed recognition score was missing in the placebo condition for one participant. The ages of two participants were missing from the Sobczak et al. (2002) data set. Additionally, from one participant, the immediate and delayed recall scores were missing in the depletion condition and from another participant, the delayed recognition in the placebo condition was missing. In the Harrison et al. (2004) study, only one immediate recall trial was presented. Therefore, the immediate recall scores of this study are not used. Finally, the study of Porter et al. (2005) missed the delayed recall and recognition score in the depletion condition and the delayed recognition score in the placebo condition of one participant.

2.7. Statistical analysis

2.7.1. Main analyses

Repeated measures analyses of variance (ANOVA) were used for statistical analysis (SPSS 12 for Windows). Within-subject factors were Treatment (TRP+ and TRP− drink) and Learning curve (for immediate recall this was Trial 1, Trial 2, and Trial 3; note again that only the first three recall sessions were implemented if one study had used more than three immediate recall trials). As between subject factors, Gender (male and female), Vulnerability (controls and participants with serotonergic vulnerability), and Age (18–40, 41–60, and 61–85 years old) were analyzed. Analyses were performed for each of the three dependent variables, namely immediate recall, delayed recall, and delayed recognition, separately. Since some of the studies used a 15-item task and the other studies a 30-item task (see Table 3), the total recall scores were recalculated into percentages to make the results comparable.

Global ANOVA was performed including Treatment, Gender, Vulnerability, and Age. Since this ANOVA revealed a significant interaction between Learning curve, Gender, and SDV ($F_{2,165} = 4.0, P = .020$) and an interaction between Treatment, Age and SDV ($F_{2,166} = 4.2, P = .016$) for the immediate recall score, two- and three-way ANOVAs were performed next for all three dependent measures. The level of significance was set at .05 throughout. Bonferroni correction was used for post hoc analyses.

2.7.2. Secondary analyses

To investigate possible association of ATD effects on memory with ATD effects on the depression scale of the POMS (mood effects), Pearson correlations were calculated. Difference values between the TRP+ and TRP− conditions were first calculated for (1) POMS depression score; (2) immediate recall; (3) delayed recall; and (4) delayed recognition accuracy. Next, correlations were calculated between the depression difference score and the memory difference scores (immediate recall, delayed recall and delayed recognition).

2.7.3. Effect size analysis

An analysis of effect sizes was performed across (1) all data, (2) the data using the 15-item task, and (3) the data using the 30-item task. Furthermore, these analyses were performed for each of the genders separately. First, the standard deviation of the distribution of differences was calculated after TRP+ and TRP−. Next, the effect size was measured, using the formula: $f = \text{difference value of delayed recall TRP+ and TRP−}/\text{distributed standard deviation}$. This value $f$ was used, together with a desired power (.8 or .9), to calculate the number of participants needed to receive significant effects. This last calculation was performed using the program GPOWER (Erdfelder et al., 1996). The results of these particular analyses can be used as power calculations for future ATD studies.

3. Results

3.1. Main analyses

3.1.1. Immediate recall

Immediate recall accuracy in percentages is summarized in Table 4 while results of significance testing can be found in Table 5. As expected, the immediate recall score increased linearly from the first until the third list presentation. TRP− treatment significantly impaired immediate recall, but no interaction between Treatment and Learning curve was found.

Gender also affected immediate recall, females recalling more words than males. A significant interaction between Treatment and Gender was found. Only females remembered significantly fewer words in the TRP− than in the TRP+ condition (Fig. 1). To find out if ATD affects immediate recall also in post-menopausal women, a separate analysis was performed for the females in the oldest age group (61–85 years). It was shown that ATD treatment also impaired immediate recall in this group ($F_{1,28} = 14.126, P = .001$).

There was an overall significant difference in immediate recall between controls and the SDV group(s). The control participants recalled more words than the vulnerable participants. Next, an analysis was performed for the SDV participants only, dividing the group into three subgroups of (1) family history for depression (unipolar and bipolar combined), (2) recovered depressed, and (3) IBS patients, in order to determine whether TRP− effects on memory differed between the vulnerability groups. The main effect of treatment on immediate recall was observed across the three vulnerability groups ($F_{1,72} = 11.091, P = .001$). Furthermore, no main Group effect was found, but an interaction between Group and Treatment ($F_{2,72} = 3.913, P = .024$). This was caused by the finding that the recovered depressed had higher immediate recall scores in the TRP+ condition than the other groups.

Age significantly affected the immediate recall score. The middle age group (41–60) had a lower recall than the other two age groups. When excluding the participants from the
Table 4
The mean percentage of recall (and standard error), and number of participants (N), during the immediate recall, delayed recall, and delayed recognition of the variables.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Placebo condition/TPR+</th>
<th>Depletion condition/TPR-</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Immediate recall</td>
<td>Delayed recall</td>
</tr>
<tr>
<td>Main treatment scores</td>
<td>58.5 (1.4), N = 180 48.4 (1.7), N = 211 85.8 (0.8), N = 184</td>
<td>54.8 (1.4), N = 180 41.6 (1.6), N = 211 83.1 (0.9), N = 184</td>
</tr>
<tr>
<td>Gender</td>
<td>Male</td>
<td>52.0 (2.0), N = 80 46.4 (2.4), N = 99 83.0 (1.2), N = 87</td>
</tr>
<tr>
<td>Female</td>
<td>63.7 (1.8), N = 100 50.2 (2.3), N = 112 88.4 (1.1), N = 97</td>
<td>57.0 (1.8), N = 100 40.6 (2.1), N = 112 85.4 (1.3), N = 97</td>
</tr>
<tr>
<td>Age</td>
<td>18–40</td>
<td>62.2 (2.1), N = 77 53.1 (2.2), N = 109 84.7 (1.2), N = 91</td>
</tr>
<tr>
<td>41–60</td>
<td>50.6 (2.8), N = 43 36.8 (3.5), N = 43 88.4 (2.0), N = 34</td>
<td>50.0 (2.8), N = 43 33.8 (3.4), N = 43 86.8 (2.2), N = 34</td>
</tr>
<tr>
<td>61–85</td>
<td>60.5 (2.4), N = 58 48.4 (3.1), N = 57 86.4 (1.5), N = 57</td>
<td>55.4 (2.4), N = 58 40.0 (3.0), N = 57 81.9 (1.7), N = 57</td>
</tr>
<tr>
<td>Vulnerability</td>
<td>Control</td>
<td>62.2 (1.8), N = 105 51.9 (2.0), N = 137 85.9 (1.0), N = 123</td>
</tr>
<tr>
<td>Vulnerable</td>
<td>53.3 (2.1), N = 75 41.9 (2.7), N = 74 85.6 (1.5), N = 61</td>
<td>49.0 (2.0), N = 75 33.8 (2.6), N = 74 82.4 (1.7), N = 61</td>
</tr>
</tbody>
</table>

Note that the immediate recall score is the score at the last immediate recall, thus recall nr. 3.

Table 5
Summary of the results of the various ANOVAs that were performed in the analysis.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Immediate recall</th>
<th>Delayed recall</th>
<th>Delayed recognition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Main effects</td>
<td>Treatment</td>
<td>F₁,₁₇₉ = 11.40, p = .001</td>
<td>F₁,₂₁₀ = 32.71, p &lt; .001</td>
</tr>
<tr>
<td></td>
<td>Recall (only immediate)</td>
<td>F₂,₁₇₈ = 496.47, p &lt; .001</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td>Gender</td>
<td>F₁,₁₇₈ = 8.86, p = .003</td>
<td>NS</td>
</tr>
<tr>
<td></td>
<td>Vulnerability</td>
<td>F₁,₁₇₈ = 13.56, p &lt; .001</td>
<td>F₁,₂₀₉ = 12.95, p &lt; .001</td>
</tr>
<tr>
<td></td>
<td>Age</td>
<td>F₁,₁₇₅ = 4.79, p = .009</td>
<td>F₂,₂₀₆ = 6.99, p = .001</td>
</tr>
<tr>
<td>Two-way interactions including the treatment factor</td>
<td>Treatment × Gender</td>
<td>F₂,₁₇₇ = 10.54, p = .001</td>
<td>F₁,₂₀₉ = 6.62, p = .011</td>
</tr>
<tr>
<td></td>
<td>Treatment × Vulnerability</td>
<td>NS</td>
<td>NS</td>
</tr>
<tr>
<td></td>
<td>Treatment × Age</td>
<td>NS</td>
<td>NS</td>
</tr>
</tbody>
</table>

Additional two-way interactions revealing significant effects

| Recall × Gender | F₂,₁₇₇ = 5.28, p = .006 | – | – |
| Recall × Vulnerability | F₂,₁₇₇ = 3.16, p = .045 | – | – |
| Recall × Age | F₄,₃₂₀ = 2.77, p = .027 | – | – |

Note: The different F-values, emerging from the different numbers of participants included in the separate analyses. All main effects and two-way interactions including the Treatment factor are shown, as well as other interactions eliciting significant results.

SDV groups, this age effect was still present, indicating that this was not due to a possible lower score by SDV participants.

Other significant effects concerning immediate recall were the interactions between Learning curve and Gender, Learning curve and SDV, and Learning curve and Age. The Gender by Learning curve interaction was due to the finding that females showed a steeper learning curve from the first till the third encoding trial. Although the males and females performed equally well at the first trial, the females outperformed the males at the third trial. The SDV × Learning curve interaction was due to the finding that the control participants showed an increased learning curve compared to the SDV group. Finally, the Age × Learning curve interaction was associated to the decreased learning curve in the middle age group compared to the other two groups.

3.1.2. Delayed recall

Delayed recall scores are listed in Table 4; results from significance testing can be seen in Table 5. Delayed recall was significantly diminished after TRP−.

As in the immediate recall condition, an interaction between Treatment and Gender was found, see Fig. 2. TRP− affected delayed recall scores of females more than those of males. Treatment also affected the delayed recall in the females of the oldest age group (F₁,₂₇ = 9.625, P = .004), who recalled fewer words after ATD.

Controls recalled more words than participants with a serotonergic vulnerability. There were no differences between the three vulnerable groups, as was found in a separate ANOVA for the three SDV groups.

Age also affected delayed recall. Post hoc analysis revealed that the youngest age group (18–40 years) recalled more words than the middle-aged group (P < .05).
3.1.3. Delayed recognition

The possible association of difference values of the depression score of the POMS (TRP+ minus TRP−) with the difference values of the last immediate recall score (TRP+ minus TRP−) was investigated using the Pearson correlation coefficient. No significant association was found ($R = .018$, $N = 118$, n.s.). This indicates that the ATD-induced immediate memory impairment is not associated with an increase in depressive feelings after the TRP− drink.

A further correlation that was examined was the correlation between the plasma TRP reduction after ATD and the difference value of the last immediate recall score (TRP+ minus TRP−). This correlation did not reach significance ($R = -.07$, $N = 163$, n.s.).

The Treatment by Gender interaction may have been due to a larger depletion of tryptophan in females than in males. Therefore, we compared the decrement in plasma TRP concentration after the TRP− drink to the TRP+ plasma between males and females. The average TRP decrement was 68.0% and did not differ between sexes ($F_{1,165} = 2.094$, n.s.).

Next, we used Pearson correlation to examine a possible relation between the difference value of the last immediate recall (TRP+ minus TRP−) and age as a continuous variable. No significant correlation was found ($R = -.017$, $N = 178$, n.s.).

3.2. Secondary analyses

Mood was not affected by treatment, as seen in the POMS depression score ($F_{1,118} = 2.866$, n.s.).

### 3.2.1. Immediate recall

The total percentage of recall is smaller than that presented in Table 4, because the recall than the men. Furthermore, the tryptophan depletion affects the both TRP recall trials, separated for males (dotted line) and females (solid line), in Fig. 1. The mean percentage of recall of the females more than of the males.

The percentage of recognition was higher after the TRP+ and the TRP− scores. The percentage of recognition was higher after Treatment significantly affected delayed recognition $\pm$ and the TRP+ condition. Note that tryptophan depletion affects performance between the difference value of the delayed recognition (TRP+ minus TRP−) than after the TRP− drink. A main effect of Gender meant that females recognized more words than males.

### 3.2.2. Delayed recall

No significant correlation was found between delayed recall and the POMS depression score ($R = -.036$, $N = 116$, n.s.), as was revealed by difference values of TRP+ minus TRP−. Furthermore, no significant correlation was found between the plasma TRP concentration and the difference values of TRP+ and TRP− ($R = -.028$, $N = 172$, n.s.). Finally, we examined the relation between the difference value of the delayed recall (TRP+ minus TRP−) and age. There was no significant relation ($R = -.027$, $N = 209$, n.s.).

### 3.2.3. Delayed recognition

ATD-induced impairment of delayed recognition scores were not associated with POMS depression difference scores ($R = .02$, $N = 118$, n.s.). Furthermore, they did not correlate with the plasma TRP concentration ($R = .024$, $N = 143$, n.s.). There was no significant correlation between the difference value of the delayed recognition score (TRP+ minus TRP−) and age ($R = .058$, $N = 182$, n.s.).

### 3.3. Effect size analysis

Table 6 shows the values of the power analysis. It can be seen that, for all dependent variables, larger effect sizes were found for the 30-item than for the 15-item list. This means that fewer participants are needed to obtain statistically significant differences between TRP+ and
TRP− conditions using a 30- compared to a 15-item word list, as can be seen from the results for powers of .8 and .9. Table 6 further shows that the delayed recall score is a more sensitive measure than immediate recall and delayed recognition, since less participants are needed to find statistically reliable results for this variable compared to the other two.

Finally, it can be seen that effect sizes are larger for females than for males, especially when using a 30-item list. This implies that less females than males are needed to reach significant results.

4. Discussion

The current ‘mega’-analysis was performed to extend the knowledge about effects of ATD on different aspects of declarative episodic memory. In brief the findings of this mega-analysis are that:

(1) Besides the repeatedly reported ATD-induced delayed memory impairments, which were confirmed in the current analysis, ATD also impaired immediate recall. However, it did not affect the steepness of the learning curve.

(2) ATD-induced memory impairment was of higher magnitude in females. This is an ATD by Gender interaction effect which is independent of other factors such as age.

(3) ATD-induced memory impairment did not differ between participants with a hypothesized vulnerability of 5-HT function.

(4) Age effects on memory were independent of those of ATD. In other words, the effects of ATD on memory were the same in age classes young, middle-aged and elderly.

(5) The plasma TRP concentration did not correlate with the magnitude of memory impairment for any of the dependent variables.

(6) The power analysis revealed that, in order to find significant results, the smallest number of participants is needed when female participants are tested using a 30-item word list.

4.1. Memory processes: ATD affects encoding or consolidation

The results of the current mega-analysis are partly in line with the view that the most likely disturbed memory
mechanism by ATD is that of memory consolidation, since indeed delayed recall was most prominently affected by ATD. However, the immediate recall effect that emerged from the mega-analysis suggests that also encoding may be impaired after ATD, albeit to a lesser extent, since it did not affect the steepness of the learning curve. Furthermore, the percentage of impairment was twice as large for delayed as for immediate recall, with reductions of 6.4% and 2.7%, respectively. The delayed recall effect may partly be due to an impaired memory encoding by ATD, but cannot entirely be explained by this phenomenon, since a larger decrement was found 15–30 min after learning than at learning itself. Encoding in this context refers to the active process of forming associations to store new information in memory, whereas consolidation refers to the forming of a permanent trace.

Izquierdo et al. (1999) proposed the existence of separate biological mechanisms to distinguish short- and long-term memory. Encoding occurs in short-term-, or working memory and the duration of unchanged short-term memory is about 30 s. According to this view, effects of ATD on immediate recall can also be explained by impaired memory consolidation, since the presentation of one 30 word list alone outlasts short-term memory. Indeed in four out of the nine studies immediate recall was affected. Furthermore, in all except one study, three or more immediate recall trials were assessed in the word list learning task. Therefore, it can be argued that ATD effects on memory consolidation could also be expressed as effects on immediate recall and therefore ATD-induced immediate recall impairments cannot only be interpreted as specific expressions of memory encoding impairments.

There are only a few imaging studies of ATD-induced memory impairments in humans (McAllister-Williams et al., 2002; Allen et al., 2006; Van der Veen et al., 2006). In the study of McAllister-Williams and colleagues, episodic memory was assessed while recording event-related potentials (ERP). They showed that, although ATD impaired delayed recall and increased ERP voltages between 500 and 1400 ms post-stimulus over posterior regions of the scalp, the ERP effect was unrelated to the neuronal correlates of retrieval. Another study by Van der Veen and co-workers manipulated depth of encoding during the presentation of a word list and could only distinguish between encoding and retrieval processes. The results showed that ATD attenuated activation in the right hippocampus in the encoding phase and that ATD did not affect brain activity in the recognition or retrieval phase. The findings of these two studies support the notion that tryptophan depletion interferes with memory encoding/consolidation processes, but does not affect memory retrieval processes.

On the basis of behavioral analysis it remains most plausible to associate the observed ATD-induced memory impairments primarily with a disturbed memory consolidation process, but this consolidation process which follows encoding, may have been influenced by prior ATD effects on memory encoding.

ATD has been shown to decrease both TRP and 5-HT levels in the rat hippocampus and additionally 5-HT levels in the striatum (Lieben et al., 2004a). Furthermore, ATD impaired memory performance in rats as measured using an object recognition task (Lieben et al., 2004b). These data appear to support a role for hippocampal 5-HT in memory processes.

Studies using ERPs have shown that posterior areas are mainly active during retrieval processes in episodic memory tasks (for a review, see Friedman and Johnson, 2000). This is in accordance with the fact that PET and fMRI studies have revealed that the medial temporal lobe, including the hippocampus, is involved in memory encoding (Köhler et al., 2000; Daselaar et al., 2003) and retrieval (Köhler et al., 2000). As noted above, ATD decreased hippocampal activity during encoding and possibly consolidation, but not retrieval processes (Van der Veen et al., 2006). These data provide more support for a role of hippocampal 5-HT specifically in memory encoding and consolidation.

4.2. Females are more sensitive to the effects of ATD

Another specific aim of this mega-analysis was to examine the effects of ATD between the sexes for the first time with an adequately powered database. In previous ATD studies in which gender effects were considered, it was shown that mood effects are larger in females than in males (Ellenbogen et al., 1996; Smith et al., 1997a; Booij et al., 2002). However, no differential effects of ATD were ever before reported on memory performance (Riedel et al., 1999; Schmitt et al., 2000), although females showed a non-significant trend towards ATD-induced memory impairment in the study of Riedel and colleagues (1999). We found that ATD effects on verbal memory are larger in females than in males on both the immediate and delayed recall scores, even though there were no significant differences between the decrements of plasma TRP after ATD in females compared to males. Irrespective of treatment, the mega-analysis further showed superior immediate and delayed recall scores of females, irrespective of treatment. This is in accordance with other studies showing that females perform better on episodic memory tasks than males (Herlitz et al., 1997, 1999; Maitland et al., 2004).

Impaired memory performance during the pre-menstrual phase of the cycle has been reported (Sherwin, 1994; Schmitt et al., 2005) and it has been shown that elevation of serotonin function by means of dietary precursor-loading with tryptophan may restore PMS related memory deficits (Schmitt et al., 2005). As explained earlier, in the current database, the influence of the menstrual cycle has effectively been controlled by not testing female participants during the pre-menstrual phase of their cycle and, furthermore, by testing a female participant within the same phase during both conditions. The present data suggest that the gender difference, i.e. the superior baseline recall performance of females as well as its perturbation by
ATD, is most pronounced during the follicular phase of the menstrual cycle. For this reason, we questioned ad hoc whether the ATD-induced recall impairment was also present in post-menopausal women in a similar way as shown for the analysis of all participants. A separate analysis for the oldest age group, consisting of 61–85 years old female participants, revealed that ATD still affects both immediate and delayed recall at this age. This result might indicate that serotonin function during the menstrual cycle does not affect memory processing and, thus, menstrual cycle cannot explain the findings in this study.

4.3. No effects of ATD on hypothesized serotonergic vulnerability

In this analysis, we pooled the data of three participant groups that are hypothesized to show an increased vulnerability for serotonergic dysfunction, namely participants with a family history for depression, recovered depressed, and mentally healthy IBS participants. We speculated that the null effect reported in previous studies was due to small group sizes. However, the present mega-analysis did not reveal any differences in effects of ATD on any of the measures. Although vulnerable participants performed worse on the verbal learning test, this did not interact with ATD effects on memory. These results suggest that participants with a presumed serotonergic vulnerability are generally not more prone to memory impairment after ATD. Possibly such vulnerability may exist for mood responses, although this was not found in the present study (effects of ATD on the depression scale of the POMS for SDV group only, $F_{1,60} = 1.118$, $P = .30$). This underscores the notion of a dissociative effect of ATD on memory and mood, which is supported by the lack of correlation between ATD-control differences scores of POMS-depression and verbal learning in the present study. However, it remains a possibility that lower dose tryptophan depletion as used by Hayward et al. (2005) may be a more sensitive way of differentiating between the effects of serotonin depletion in vulnerable and non-vulnerable groups.

4.4. Effects of ATD and age on memory are independent

Aging alters cholinergic and serotonergic hippocampal innervations and the release of both acetylcholine and serotonin in rats (Birthelmer et al., 2003). The hippocampal changes in rats do not interact with memory processing, since both aged rats with good and poor performance in the spatial water-maze task have altered hippocampal innervations (Birthelmer et al., 2003). This might suggest that possible age-related changes in memory performance are not due to 5-HT variations.

In humans, a decline in serotonin function with increasing age has been found (Van Dyck et al., 2000; Baeken et al., 1998). It was shown that the effect of at least one specific 5HT<sub>2A</sub> receptor polymorphism on episodic memory is age-dependent (Papassotiropoulos et al., 2005), thus suggesting that a serotonergic vulnerability alters with age. Furthermore, ATD may impair some aspects of cognitive function, amongst which memory processing, to a greater extent in elderly participants suffering from Alzheimer’s disease than in age-matched controls (Porter et al., 2000, 2003b). This has been attributed to an imbalance between serotonergic and cholinergic systems in the brain (Garcia-Alloza et al., 2005) and indicates that an induced 5-HT deficit may exacerbate the effect of a pre-existing cognitive deficit, possibly associated with cholinergic dysfunction, as seen in Alzheimer’s disease (Porter et al., 2000, 2003b).

However, to the best of our knowledge, age-related effects of ATD have, not yet been studied to date. In the present mega-analysis, we found a main effect of age on memory performance indicative of age-related decline but we did not find an Age by ATD interaction effect. The absence of an interaction in this mega-analysis does not support the notion that age-related memory decline is associated with an increased serotonergic vulnerability. It must be kept in mind, however, that the methodology of the verbal learning task may have caused this effect. The studies, in which young participants were used, mostly employed the 30 word version of the verbal learning task. In the studies that used older participants, the 15 word version was always used. Even though we corrected the scores on both versions to be able to interpret our current results, it may be the case that the 15 word verbal learning test is easier to perform. This may have obscured a difference and may have rendered the older participants to ‘outperform’ themselves compared to the younger participants.

4.5. Effect sizes

The current analyses revealed that statistically reliable differences between TRP<sup>+</sup> and TRP<sup>−</sup> conditions can be obtained with fewer participants using the 30-item word lists than when using 15-item lists. Furthermore, effect sizes were larger in females than in males, which indicated that fewer participants are needed if only females participate in a study.

One of the differences between a 30- and a 15-item word list is the chance of reaching a ceiling effect. Since fewer items are presented in the 15-item list, the possibility of remembering all words is larger compared to the 30-item list. Furthermore, in most of the studies using 15 items (Porter et al., 2000, 2005; Hughes et al., 2003), five immediate recall trials were used, whereas the studies that used 30-item lists performed three immediate recall trials (Riedel et al., 1999; Schmitt et al., 2000; Sobczak et al., 2002; Kilkens et al., 2004). This methodological difference increases the occurrence of a ceiling effect even more.

When comparing the results of the separate studies (see Table 3), it can be seen that treatment effects were mainly obtained in the studies using 30 items. This underscores the
fact that the 30-item lists are more sensitive for finding ATD effects than the 15-item lists.

It must be noted that ATD impaired delayed recognition. However, when reviewing the results from Table 6, it seems that delayed recognition is not a sensitive enough measure, since a high number of participants are needed in order to obtain a difference. First of all, ceiling effects are even more of a problem at recognition than at recall. This causes the effect sizes to be rather small. Additionally, it has been suggested that the increase in number of false alarms after ATD reveals the treatment effect better than does the number of correctly accepted and rejected items (Meeter et al., 2006). Schmitt and colleagues (2000) did take the number of false alarms into account and calculated sensitivity measures for the delayed recognition. They found that this measure was impaired after ATD, whereas other studies only including the number of correctly accepted and rejected items failed to obtain such an effect (Hughes et al., 2003; Hayward et al., 2005). It is, therefore, suggested that false alarm scores should be evaluated after ATD as well.

4.6. Conclusion

In conclusion, the present mega-analysis has shown that not only delayed recall, but to a lesser extent also immediate recall is affected by ATD. These results are most likely explained by an effect of lowering 5-HT on a memory consolidation process, but this consolidation process which follows encoding, may have been influenced by prior ATD effects on memory encoding. Future research may be specifically aimed at unraveling these processes. In addition, neither a hypothesized serotonergic dysregulation and/or vulnerability, nor age, appear to have an impact on ATD-induced memory impairments. ATD, however, does affect females more than males, as indicated by a greater decrement in both immediate and delayed recall scores. Therefore, the only factor that actually has the properties of a serotonergic vulnerability factor in memory performance is Gender. This is the first report to establish that female gender is associated with an increased sensitivity for memory impairment due to low serotonin.

Acknowledgements

We would like to thank the following people for their help in data collection and analysis of the original papers: Jos Adam, Clive Ballard, Linda Booij, Robert-Jan Brummer, Graham Burrows, Phil Cowen, Nikolaas Deutz, Peter Gallagher, Guy Goodwin, John Gray, Ben Harrison, Gail Hayward, Eef Hogervorst, Adriaan Honig, Brenda Jorissen, Tom Kelly, Tineke Klaassen, Albert Leentjens, Brian Lunn, David Matthews, Trevor Norman, John O’Brien, James Olver, Andrew Phipps, Jos Prickaerts, Marije aan het Rot, Ailsa Scott, Pamela Stevenson, Mary Stewart, Martin van Boxtel, Michiel van Nieuwenhoven, Herman van Praag, Astrid van Someren, Frans Verhey, Lucy Walker, Keith Wesnes, and Allan Young.

References


