The effects of acute tryptophan depletion and serotonin transporter polymorphism on emotional processing in memory and attention

Jonathan P. Roiser1,2, Ulrich Müller3, Luke Clark3 and Barbara J. Sahakian1

1 Department of Psychiatry, Box 189, Addenbrooke’s Hospital, Cambridge, UK
2 Institute of Neurology, Queen Square House, London, UK
3 Department of Experimental Psychology, Cambridge, UK

Abstract
Polymorphism at the serotonin transporter linked polymorphic region (5-HTTLPR) has been associated with neuroticism, increased risk for affective disorders and greater vulnerability to mood change following serotonin (5-HT) depletion. The aim of the present study was to investigate whether the cognitive effects of 5-HT depletion were differentially affected by genotype at the 5-HTTLPR polymorphism, using neuropsychological measures of memory and attention. We utilized the acute tryptophan depletion (ATD) technique to temporarily reduce 5-HT synthesis in two groups of healthy volunteers pre-selected on the basis of 5-HTTLPR genotype, 15 of the ll genotype and 15 of the ss genotype, in a double-blind, placebo-controlled crossover design. As expected, ATD resulted in a robust reduction in plasma tryptophan concentration in both genotype groups. However, the genotype groups differed in terms of the effect of ATD on cognitive performance. The ss genotype group showed impaired verbal recall following depletion, while episodic memory was unimpaired by ATD in the ll genotype group. Averaging across depletion condition, the ss genotype group outperformed the ll genotype group on tests of episodic memory and attention. Neither group was significantly affected by ATD on measures of emotional state. These data confirm previous reports that ss individuals are particularly vulnerable to 5-HT depletion, but extend these findings to the cognitive domain. The unexpected finding that ss volunteers showed improved memory and attention relative to ll volunteers suggests a possible evolutionary advantage to possession of the s allele, which may offset the disadvantage of vulnerability to depression following stressful life events.

Received 3 February 2006; Reviewed 29 March 2006; Revised 30 May 2006; Accepted 31 May 2006; First published online 8 August 2006

Key words: Acute tryptophan depletion (ATD), attention, emotion, memory, serotonin transporter gene-linked polymorphic region (5-HTTLPR).

Introduction
The technique of acute tryptophan depletion (ATD) has been employed in a variety of studies in order to understand the role of serotonin (5-HT) in various cognitive functions and neuropsychiatric disorders. ATD has been shown to induce a transient return of symptoms in patients recovered from depression, both in unmedicated patients and those stabilized on selective serotonin reuptake inhibitors (e.g. Delgado et al., 1990; Leyton et al., 1997; Neumeister et al., 2004). The effects of ATD have also been examined in healthy volunteers, with most studies reporting that ATD-induced mood change in healthy volunteers is relatively uncommon (Riedel, 2004, although see Young et al., 1985; Smith et al., 1987).

The findings of studies investigating the effect of ATD on neuropsychological performance in healthy volunteers are fairly consistent. Episodic memory is commonly impaired by ATD (McAllister-Williams et al., 2002; Riedel et al., 1999; Rubinsztein et al., 2001), consistent with the finding that 5-HT is reduced in Alzheimer’s disease (Meltzer et al., 1998), the primary
feature of which is episodic memory impairment. Tasks with an emotional or reward-related component are also affected by ATD; previous studies have reported reduced sensitivity to reward and a negative emotional bias in healthy volunteers following ATD (Klaassen et al., 2002; Murphy et al., 2002; Rogers et al., 2003; Rubinsztein et al., 2001). By contrast, tests designed to measure ‘executive function’, for example attention, working memory, planning and cognitive flexibility, are normally unaffected (Murphy et al., 2002; Park et al., 1994) or even improved (Gallagher et al., 2003; Schmitt et al., 2000) following ATD.

It is clear that there is large individual variation in response to ATD (Booij et al., 2003), which may be related to genetic factors (Crean et al., 2002; LeMarquand et al., 1999; Neumeister et al., 2002; Quintin et al., 2001). For example, one study to date has reported that women homozygous for the s allele at the 5-HT transporter linked polymorphic region (5-HTTLPR) were vulnerable to mood change following ATD, while those homozygous for the l allele did not show mood change (Neumeister et al., 2002). In heterozygous participants, those with a family history of depression showed greater mood change than those with no family history of depression. However, no published study has yet examined the effect of 5-HTTLPR polymorphism on response to ATD in terms of cognitive performance.

The aim of this study was to investigate the effect of ATD on mood and neuropsychological performance in healthy volunteers of ll or ss genotypes at the 5-HTTLPR. To ensure equal numbers of each genotype, we pre-selected volunteers by genotype and employed a double-blind, placebo-controlled cross-over design. We chose to test volunteers homozygous either for the l or s allele at the 5-HTTLPR in order to maximize any differences between genotype groups. Participants were administered a neuropsychological assessment based on the cognitive sequelae of depression, including memory and attention. It was predicted that individuals carrying the ss genotype would show greater vulnerability to ATD than ll individuals, as indexed by measures of mood, emotional processing and episodic memory, but not attention.

Method

Participants and genetic screening (Phase I)

This study was carried out in two phases, each with its own informed consent process. Initially, volunteers were recruited by advertisement in the community. All participants who responded to the advertisement were invited to attend an initial screening session (Phase I) where they were administered a general health questionnaire, including a section on psychiatric disorders, the Adult Impulsiveness, Venturesomeness and Empathy Scale (IVE; Eysenck and Eysenck, 1991), the Barratt Impulsiveness Scale (BIS-11 Patton et al., 1995) and the Beck Depression Inventory (BDI; Beck et al., 1961). A 10-ml blood sample was taken, from which DNA was extracted in order to determine 5-HTTLPR genotype. All participants completing Phase I were compensated £7 for their time and travel expenses. The study was approved by the Cambridge Research Ethics Committee (Project no. 03/266), and all participants gave informed consent, including a section consenting to the storage of their DNA and contact details. Inclusion criteria for the tryptophan depletion study (Phase II) were: ll or ss genotype at the 5-HTTLPR. Exclusion criteria were: any history of Axis I psychiatric disorders, including any treatment for depression or anxiety disorders (including a score >9 on the BDI at screening); any history of neurological illness or closed head injury; any history of drug or alcohol dependence. Approximately 50 volunteers who attended the screening session satisfied the inclusion and exclusion criteria and were invited to take part in Phase II. Seven of the participants (3 ll, 4 ss) completing the ATD protocol were included as controls in a previous study (Roiser et al., 2005).

Experimental procedure – ATD (Phase II)

All participants were informed of the details and procedure of Phase II either by telephone or at the screening session. All participants attending Phase II gave informed consent, were compensated £35 for their time and travel expenses for each testing day completed, and the study was approved by the Cambridge Research Ethics Committee (Project no. 03/338). All volunteers in Phase II were Caucasian, in order to avoid confounds due to population stratification. Participants arrived at the Wellcome Trust Clinical Research Facility at the Addenbrooke’s Centre for Clinical Investigation at approximately 08:30 hours having fasted from 24:00 hours the previous night, drinking only water. Following baseline (T₀) measurements of mood and a 10-ml blood sample, participants were administered either a balanced amino-acid drink (Trp +) or the same mixture with the absence of tryptophan (Trp−) (for constituents see below). The amino acids were dissolved in ~300 ml tap water with fruit flavouring to improve palatability.

Participants and genetic screening (Phase I)

This study was carried out in two phases, each with its own informed consent process. Initially, volunteers were recruited by advertisement in the community. All participants who responded to the advertisement were invited to attend an initial screening session (Phase I) where they were administered a general health questionnaire, including a section on psychiatric disorders, the Adult Impulsiveness, Venturesomeness and Empathy Scale (IVE; Eysenck and Eysenck, 1991), the Barratt Impulsiveness Scale (BIS-11 Patton et al., 1995) and the Beck Depression Inventory (BDI; Beck et al., 1961). A 10-ml blood sample was taken, from which DNA was extracted in order to determine 5-HTTLPR genotype. All participants completing Phase I were compensated £7 for their time and travel expenses. The study was approved by the Cambridge Research Ethics Committee (Project no. 03/266), and all participants gave informed consent, including a section consenting to the storage of their DNA and contact details. Inclusion criteria for the tryptophan depletion study (Phase II) were: ll or ss genotype at the 5-HTTLPR. Exclusion criteria were: any history of Axis I psychiatric disorders, including any treatment for depression or anxiety disorders (including a score >9 on the BDI at screening); any history of neurological illness or closed head injury; any history of drug or alcohol dependence. Approximately 50 volunteers who attended the screening session satisfied the inclusion and exclusion criteria and were invited to take part in Phase II. Seven of the participants (3 ll, 4 ss) completing the ATD protocol were included as controls in a previous study (Roiser et al., 2005).

Experimental procedure – ATD (Phase II)

All participants were informed of the details and procedure of Phase II either by telephone or at the screening session. All participants attending Phase II gave informed consent, were compensated £35 for their time and travel expenses for each testing day completed, and the study was approved by the Cambridge Research Ethics Committee (Project no. 03/338). All volunteers in Phase II were Caucasian, in order to avoid confounds due to population stratification. Participants arrived at the Wellcome Trust Clinical Research Facility at the Addenbrooke’s Centre for Clinical Investigation at approximately 08:30 hours having fasted from 24:00 hours the previous night, drinking only water. Following baseline (T₀) measurements of mood and a 10-ml blood sample, participants were administered either a balanced amino-acid drink (Trp +) or the same mixture with the absence of tryptophan (Trp−) (for constituents see below). The amino acids were dissolved in ~300 ml tap water with fruit flavouring to improve palatability.
Participants attended two testing sessions, separated by at least 1 wk in order to avoid carry-over effects. Both participants and researcher were blind as to the nature of the drinks. For the next 5 h, participants rested at the research centre, and were allowed water ad libitum as well as decaffeinated tea or coffee (without milk). At 12:00 hours participants were given a low-protein lunch of salad, crackers, jam and either an apple or an orange. At 14:00 hours (T5), a second 10-ml blood sample was taken, participants' mood was assessed and tests of memory and attention were administered. Thirty volunteers completed Phase II, 15 of the ll genotype (8 male, 7 female) and 15 of the ss genotype (9 male, 6 female). Female participants were tested in the follicular phase of their menstrual cycle.

Amino-acid mixtures

The quantities of amino acids in each drink were based on those used by Sobczak et al. (2002), who reported a substantial reduction in plasma tryptophan concentration. Amino-acid mixtures were as follows:

Trp +: t-alanine, 4.1 g; l-arginine, 3.7 g; l-cystine, 2.0 g; glycine, 2.4 g; l-histidine, 2.4 g; l-isoleucine, 6 g; l-leucine, 10.1 g; l-lysine, 6.7 g; l-methionine, 2.3 g; l-proline, 9.2 g; l-phenylalanine, 4.3 g; l-serine, 5.2 g; l-threonine, 4.9 g; l-tyrosine, 5.2 g; l-valine, 6.7 g and l-tryptophan, 3.0 g; total, 78.0 g.

Trp −: t-alanine, 4.1 g; l-arginine, 3.7 g; l-cystine, 2.0 g; glycine, 2.4 g; l-histidine, 2.4 g; l-isoleucine, 6 g; l-leucine, 10.1 g; l-lysine, 6.7 g; l-methionine, 2.3 g; l-proline, 9.2 g; l-phenylalanine, 4.3 g; l-serine, 5.2 g; l-threonine, 4.9 g; l-tyrosine, 5.2 g and l-valine, 6.7 g; total, 75.0 g.

For female participants, the same ratios of amino acids were used, but with a 20% reduction in quantity to take into account lower bodyweight.

Biochemical measures

Plasma was separated by centrifugation and stored at −20 °C. Plasma total amino-acid concentrations (tyrosine, valine, phenylalanine, isoleucine, leucine and tryptophan) were measured by means of HPLC with fluorescence end-point detection and pre-column sample derivatization adapted from the methods of Furst et al. (1990). Norvaline was used as an internal standard. The limit of detection was 5 nmol/ml using a 10-μl sample volume, and inter- and intra-assay coefficients of variation were <15% and <10% respectively.

Mood rating scales

At T0 and T5 participants were administered the following mood assessments:

Modified Hamilton Depression Rating Scale (HDRS; Hamilton, 1960)

The HDRS is an observer rating scale for depressive symptomology, conducted in an interview format. For each question, the interviewer scores the patient’s report with higher numbers indicating increasing severity. This version of the HDRS excluded the five symptoms that could not change over the course of a study day (insomnia initial, insomnia middle, insomnia late, genital symptoms, loss of weight). Therefore, the number of questions was reduced to 16, and the maximum possible score to 49 (Murphy et al., 2002).

Positive and Negative Affect Scale (PANAS; Watson et al., 1988)

The PANAS consists of two 10-item mood scales designed to assess positive and negative experience. Participants are presented with a list of 20 affect-related descriptors and asked to respond on a scale from 1 (‘very slightly or not at all’) to 5 (‘extremely’). The positive-affect descriptors included: interested, alert, excited, inspired, strong, determined, attentive, enthusiastic, active, and proud. The negative affect descriptors included: irritable, distressed, ashamed, upset, nervous, guilty, scared, hostile, jittery, and afraid.

Cognitive assessment

In order to assess cognitive function following ATD, participants were administered a series of computerized tests of memory and attention: an emotionally loaded version of the Directed Forgetting paradigm (Ullsperger et al., 2000; Zacks et al., 1996), the CANTAB Pattern Recognition Memory (PRM) test (Sahakian et al., 1988) and the CANTAB Affective Go/No-go (AGNG) test (Murphy et al., 1999). A brief description of each of the tests is provided below. Participants also performed a test of incentive motivation and the Stop Signal Reaction Time test, although results for these test have been included in separate communications (Clark et al., 2005; Roiser et al., 2006). All participants sat ~60 cm from a touch-sensitive computer screen controlled by an Advantech Pentium personal computer (Model PPC-120T-RT). For all tests where the order of stages, blocks, or trials could vary,
stage, block or trial order was counterbalanced across drink order.

Affective Directed Forgetting (ADF; Ullsperger et al., 2000)

This test is based on a well-established paradigm that probes the ability of an individual to selectively inhibit long-term encoding/retrieval of a stimulus (Zacks et al., 1996). There are two stages to the test. First, participants carry out an encoding phase, during which they are instructed to remember some material [termed ‘to be remembered’ (TBR)] and forget other material [termed ‘to be forgotten’ (TBF)]. Following this, there is a retrieval phase, during which participants must attempt to remember all material from the encoding phase, regardless of the instruction received at the time (MacLeod, 1999). We employed an item-cued version of the Directed Forgetting test, as opposed to a list version, to minimize practice effects.

In the encoding phase, on each trial a fixation point is presented for 200 ms, followed by a 2000 ms blank screen, then a word for 250 ms, followed by a 2500 ms blank screen, and then either the letter ‘R’ or the letter ‘F’ for 400 ms, followed by a 200 ms blank screen. Participants are instructed that they must try to remember words followed by the letter R (TBR words), and forget words followed by the letter F (TBF words). The 2500-ms gap is included to ensure that participants attend to the word initially and engage in some kind of encoding (and not divert their attention). In this version, positive, negative and neutral words were used, chosen from the Affective Norms for English Words (Bradley and Lang, 1999). Positive, negative and neutral sets of words were matched for frequency, length and number of syllables. Four blocks of 36 words are presented, and at the end of each block participants are required to write down as many TBR words as they can (recall test). In each block, 10 positive words, 10 negative words and 10 neutral words (5 TBR and 5 TBF of each) are presented, in a pseudo-random order. In addition, six ‘buffer’ words are used in each block, all neutral TBR items, making up trials 1, 2, 3 and 34, 35, 36. The buffer words are included in order to minimize primacy and recency effects.

Following a 20-min delay, during which participants completed other tests, participants carry out a surprise recognition test. On each trial, a fixation point appears for 1000 ms, a word is then presented for 250 ms and then the screen is blank for up to 2000 ms. As soon as each word is presented, participants must decide whether it is an ‘old’ word (one that they recognize from the previous four blocks) or a ‘new’ word (one that did not appear in the previous four blocks), and respond appropriately using one of two buttons. Participants are explicitly told that if they think they have seen the word in the previous four blocks they should press the ‘old’ button, regardless of whether they believe the word was TBR or TBF in the encoding phase. In total, 240 words are presented, 120 old words and 120 new words (the buffer words are not presented in this stage). Participants must respond within 2250 ms, and the next trial starts immediately following their response. Halfway through the recognition phase, participants are permitted a short break.

Previous studies have found that TBR words are both recalled and recognized significantly more accurately than TBF words, which has been interpreted as reflecting the ability to prevent formation of a memory if instructed to do so (MacLeod, 1999; Paz-Caballero et al., 2004). However, it is also possible that this effect may be due to increased rehearsal of TBR items (Lehman et al., 2001). Measures arising from this test are percent correct (TBR) and intrusions (TBF) in the recall stage, and percent correct and response latency at the recognition stage for positive, negative and neutral words.

CANTAB Pattern Recognition Memory (PRM; Sahakian et al., 1989) (see www.camcog.com)

In this memory test, participants are shown a series of 12 abstract patterns and are instructed to remember them. Following a 5-s delay, each pattern is then shown again to the participant in reverse order paired with a novel pattern. Participants are required to make a forced-choice discrimination by touching the pattern they have seen previously. Feedback is provided to the participant by way of green ticks and red crosses. This procedure is then repeated with a further 12 patterns. Following a delay of ~20 min, the recognition phase of the task is repeated with the same forced-choice trials. Measures arising from this test are percent correct and latency.

Affective Go/No-go (AGNG; Murphy et al., 1999) (see www.camcog.com)

In the AGNG test, emotionally toned words are presented in the middle of the screen sequentially. Half the words are targets, to which the participant is required to respond, while half are distractors, which the participant is instructed to ignore. Participants respond by pressing the space bar as quickly as possible.
Words are presented for 300 ms, with an inter-stimulus interval of 900 ms. A 500-ms 450-Hz tone sounds for each commission error, but not for omission errors. The task consists of 10 blocks of 18 stimuli – nine positive or ‘happy’ (H) and nine negative or ‘sad’ (S) words per block. The first two blocks are treated as practice blocks. In each block, either positive or negative words are specified as targets, with targets for the 10 blocks presented in a HHSSHHSSHH or SSHHSSHHSS order. Four blocks can therefore be considered ‘shift’ blocks, where participants must start responding to stimuli that were previously distractors, while withholding response to stimuli that were previously targets. Measures arising from this task are correct response latency, number of commission errors (false alarms) and number of omission errors (misses) on each block.

Genetic analysis
Genotyping was performed as described previously (Furlong et al., 1998).

Statistical analysis
Data were analysed using SPSS version 12 (SPSS Inc., Chicago, IL, USA). This study employed a two-group, within-subjects, placebo-controlled crossover design, with repeated measures on various indices of cognitive function (e.g. stage of test, emotion of stimulus). Therefore, mixed ANOVA was employed where test assumptions were met (i.e. if variances were equivalent and data were normally distributed). In general, treatment (Trp+/Trp−) and stage of test were entered as within-subjects factors, and genotype group (ll or ss) was entered as the between-subjects factor. Drink order (Trp−/Trp+ or Trp+/Trp−) and battery order (forward/reverse) were counterbalanced, but were entered as additional between-subjects factors. If the main effect of drink order/battery order and interactions of these factors with treatment and/or genotype were found to be non-significant, data were collapsed across drink order/battery order for subsequent analyses. Post-hoc analyses were carried out by constructing appropriate ANOVAs for each comparison of interest. Where there was a significant treatment × order interaction in the omnibus ANOVA, post-hoc analyses examining effects of treatment included order as a between-subjects factor. In cases where there was a departure from the assumption of homogeneity of covariance in the repeated-measures ANOVA, an epsilon (ε) factor was calculated and used to adjust degrees of freedom accordingly, using the Huynh–Felt procedure (Howell, 2002). Where appropriate, proportion data were arcsine-transformed and latency data log-transformed prior to analysis to reduce skew and stabilize variances (Howell, 2002), although untransformed data are presented for clarity.

Results

Demographic and questionnaire measures
Demographic and questionnaire measures are reported in Table 1. The genotype groups did not
differ significantly on any demographic or personality measure, or in terms of diagnosis of depression in a first-degree relative.

Plasma amino-acid concentrations

Due to difficulties with taking blood, plasma samples were not available for three participants at all four time-points (one ll genotype, two ss genotype). For plasma total tryptophan concentration an expected significant treatment × time interaction was evident \([F(1, 25) = 102.0, p < 0.001]\). Total plasma tryptophan concentration increased from \(T_0\) to \(T_5\) by 65% when Trp was administered \([F(1, 25) = 49.2, p < 0.001]\), but fell from \(T_0\) to \(T_5\) by 69% when Trp was administered \([F(1, 25) = 52.1, p < 0.001]\). This interaction was equivalent between the two genotype groups \([treatment \times time \times genotype interaction: F(1, 25) < 1]\) (see Table 2). Plasma tryptophan concentrations and tryptophan: S LNAA ratio were unaffected by gender following either Trp + or Trp − days for the il genotype group \([p > 0.1]\), but the ss genotype group showed a significant treatment × time × genotype interaction \([F(1, 25) = 4.12, p = 0.012]\). While in the il genotype group \([F(1, 29) = 7.7, p = 0.013]\), the effect of time on depressive symptoms (as measured by the modified HDRS) was dependent on genotype \([time \times genotype interaction: F(1, 29) = 24.1, p < 0.001]\). The effect of treatment was dependent on genotype \([treatment \times time \times genotype interaction: F(1, 29) = 5.9, p < 0.01]\) (see Table 2). Plasma total tryptophan concentration was equivalent between the two genotype groups \([treatment \times time \times genotype interaction: F(1, 25) = 1.1, p = 0.3]\). A similar treatment × time × genotype interaction \([F(1, 25) = 11.4, p < 0.001]\) was present for the tryptophan: S LNAA (large neutral amino acid) ratio \([F(1, 25) = 15.4, p < 0.001]\), which was not significantly different from the trp − day interaction \([F(1, 25) = 1.1, p > 0.1]\).

Table 2. Amino acid concentrations and mood rating scales. Values represent mean (S.D.)

<table>
<thead>
<tr>
<th>Test</th>
<th>Measure</th>
<th>Placebo</th>
<th>Depleted</th>
<th>Placebo</th>
<th>Depleted</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>il genotype</td>
<td>(T_0)</td>
<td>(T_5)</td>
<td>(T_0)</td>
<td>(T_5)</td>
</tr>
<tr>
<td>Plasma tryptophan</td>
<td>Total tryptophan (μmol/l)</td>
<td>43.3 (20.1)</td>
<td>77.9 (51.6)</td>
<td>49.4 (13.6)</td>
<td>19.1 (14.4)</td>
</tr>
<tr>
<td></td>
<td>Trp: S LNAA ratio</td>
<td>0.10 (0.06)</td>
<td>0.13 (0.06)</td>
<td>0.11 (0.04)</td>
<td>0.04 (0.03)</td>
</tr>
<tr>
<td>PANAS</td>
<td>Positive</td>
<td>30.0 (8.2)</td>
<td>24.8 (5.3)</td>
<td>30.4 (6.4)</td>
<td>27.6 (7.6)</td>
</tr>
<tr>
<td></td>
<td>Negative</td>
<td>12.1 (3.2)</td>
<td>11.1 (2.3)</td>
<td>11.1 (1.4)</td>
<td>10.9 (1.2)</td>
</tr>
<tr>
<td></td>
<td>Total score</td>
<td>1.9 (2.8)</td>
<td>1.3 (1.2)</td>
<td>1.4 (1.7)</td>
<td>0.8 (1.1)</td>
</tr>
<tr>
<td>HDRS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

HDRE, Hamilton Depression Rating Scale; PANAS, Positive and Negative Affect Scale; Trp: S LNAA, tryptophan: S large neutral amino acids.
trend towards lower positive scores on the Trp− day than on the Trp+ day \(F(1, 28)=6.4, p=0.017\). The \(SS\) genotype group showed significantly worse recall following Trp− compared to Trp+ \(F(1, 14)=4.9, p=0.042\), while the \(LL\) genotype group showed no difference in recall between Trp− and Trp+ \(F(1, 14)=1.0, p=0.35\). Furthermore, the \(SS\) genotype group showed a trend towards recalling more words than the \(LL\) genotype group following Trp+ \(F(1, 28)=3.8, p=0.062\), but not following Trp− \(F(1, 28)<1\) (see Figure 1). Recall accuracy was not affected by emotional valence, and all other interactions and main effects were non-significant. Rates of intrusion errors at the recall stage were very low (mean <3%), with many participants making no intrusion errors, and these data were therefore unsuitable for parametric statistical analysis.

For the recognition phase, data from one \(LL\) female participant were missing due to computer failure, and data from one male \(LL\) participant were excluded due to a failure to follow test instructions. Therefore, the following analyses are based on data from 28 participants.

The effect of instruction at encoding (‘Remember’ or ‘Forget’) on recognition accuracy was dependent on emotional valence \(F(2, 52)=4.4, p=0.014\). The ‘directed forgetting effect’ (TBR minus TBF) was significantly greater for neutral words than for happy words \(F(1, 26)=8.5, p=0.007\) and sad words \(F(1, 26)=8.1, p=0.008\) and sad TBF words \(F(1, 26)=12.7, p=0.001\), while recognition of TBR words was unaffected by emotional valence \(F(2, 52)<1\). This suggests that, regardless of treatment and genotype, participants were more successful at inhibiting encoding for neutral words than for emotional words (see Figure 2). As expected, participants recognized more TBR words than TBF words \(F(1, 26)=62.7, p<0.001\), but recognition accuracy was unaffected by treatment or genotype, and all other interactions were non-significant.

Analysis of latency data revealed that participants responded more quickly to TBR words than TBF words \(F(1, 26)=5.1, p=0.032\), and showed a strong trend towards responding more quickly to happy words than sad words \(F(2, 52)=3.1, p=0.053\). However, latency was unaffected by treatment or genotype and all interactions were non-significant.

**Affective Directed Forgetting (ADF)**

Analysis of accuracy data at the recall stage revealed a significant treatment × genotype interaction \(F(1, 28)=6.4, p=0.017\). The \(SS\) genotype group showed significantly worse recall following Trp− compared to Trp+ \(F(1, 14)=5.0, p=0.042\), while the \(LL\) genotype group showed no difference in recall between Trp− and Trp+ \(F(1, 14)=2.0, p=0.18\). Furthermore, the \(SS\) genotype group showed a trend towards recalling more words than the \(LL\) genotype group following Trp+ \(F(1, 28)=3.8, p=0.062\), but not following Trp− \(F(1, 28)<1\) (see Figure 1). Recall accuracy was not affected by emotional valence, and all other interactions and main effects were non-significant. Rates of intrusion errors at the recall stage were very low (mean <3%), with many participants making no intrusion errors, and these data were therefore unsuitable for parametric statistical analysis.

For the recognition phase, data from one \(LL\) female participant were missing due to computer failure, and data from one male \(LL\) participant were excluded due to a failure to follow test instructions. Therefore, the following analyses are based on data from 28 participants.

The effect of instruction at encoding (‘Remember’ or ‘Forget’) on recognition accuracy was dependent on emotional valence \(F(2, 52)=4.4, p=0.014\). The ‘directed forgetting effect’ (TBR minus TBF) was significantly greater for neutral words than for happy words \(F(1, 26)=8.4, p=0.008\) and sad words \(F(1, 26)=8.5, p=0.007\). This decrease in the directed forgetting effect for emotional words was entirely due to the effect on TBF words; recognition of neutral TBF words was worse than recognition of happy TBF words \(F(1, 26)=8.5, p=0.007\) and sad TBF words \(F(1, 26)=12.7, p=0.001\), while recognition of TBR words was unaffected by emotional valence \(F(2, 52)<1\). This suggests that, regardless of treatment and genotype, participants were more successful at inhibiting encoding for neutral words than for emotional words (see Figure 2). As expected, participants recognized more TBR words than TBF words \(F(1, 26)=62.7, p<0.001\), but recognition accuracy was unaffected by treatment or genotype, and all other interactions were non-significant.

Analysis of latency data revealed that participants responded more quickly to TBR words than TBF words \(F(1, 26)=5.1, p=0.032\), and showed a strong trend towards responding more quickly to happy words than sad words \(F(2, 52)=3.1, p=0.053\). However, latency was unaffected by treatment or genotype and all interactions were non-significant.

**Pattern Recognition Memory**

Data from 1 female \(SS\) participant were lost due to computer failure. Analysis of accuracy data revealed that the \(SS\) genotype group was more accurate than the \(LL\) genotype group overall \(F(1, 27)=5.1, p=0.032\), and that this difference in accuracy was particularly pronounced at the delayed stage of the test \(genotype\times delay interaction: F(1, 27)=4.6, p=0.042; see Figure 3\). As expected, participants were more accurate at the immediate stage than the delayed stage of the test \(F(1, 27)=26.6, p<0.001\), but accuracy was unaffected by treatment and all other interactions were non-significant. All main effects and interactions for PRM latency were non-significant.

**Affective Go/No-go (AGNG)**

Analysis of latency data revealed no evidence of a bias towards sad words following Trp−, nor was the effect of Trp− on response to emotional words dependent on genotype; all main effects and interactions for latency were non-significant. Analysis of commission error data revealed a significant four-way interaction between treatment, emotional valence, shift condition and genotype \(F(1, 28)=4.4, p=0.045\). However, post-hoc analysis could not clarify this result (all simple effects \(p>0.1\)). The main effects of treatment, valence, shift condition and genotype on commission errors, and all other interactions, were non-significant. The \(LL\) genotype group made significantly more omission errors overall than the \(SS\) genotype group \(F(1, 28)=4.4, p=0.045\) (see Figure 4). In addition, participants made significantly more omission errors when targets were
happy than when targets were sad \( F(1, 28) = 4.9, \ p = 0.035 \). However, omission errors were unaffected by treatment or shift condition, and all interactions were non-significant.

TABLE 3. Neuropsychological data. Values represent mean (S.D.)

<table>
<thead>
<tr>
<th>Test</th>
<th>Measure</th>
<th>ll genotype Placebo</th>
<th>Depleted</th>
<th>ss genotype Placebo</th>
<th>Depleted</th>
</tr>
</thead>
<tbody>
<tr>
<td>Directed forgetting – recall</td>
<td>Percent correct</td>
<td>31.3 (18.9)</td>
<td>35.0 (19.3)</td>
<td>44.3 (16.2)</td>
<td>40.3 (15.5)</td>
</tr>
<tr>
<td></td>
<td>Happy</td>
<td>36.0 (23.2)</td>
<td>33.0 (23.4)</td>
<td>47.0 (22.4)</td>
<td>40.6 (21.1)</td>
</tr>
<tr>
<td></td>
<td>Neutral</td>
<td>30.3 (23.9)</td>
<td>38.3 (20.3)</td>
<td>46.6 (21.6)</td>
<td>43.6 (22.0)</td>
</tr>
<tr>
<td>Directed forgetting –</td>
<td>TBR percent correct</td>
<td>65.0 (21.4)</td>
<td>59.2 (15.0)</td>
<td>73.9 (17.6)</td>
<td>68.5 (13.5)</td>
</tr>
<tr>
<td>recognition</td>
<td>TBR latency</td>
<td>703.3 (160.3)</td>
<td>644.7 (145.5)</td>
<td>653.9 (97.5)</td>
<td>706.4 (193.4)</td>
</tr>
<tr>
<td></td>
<td>TBF percent correct</td>
<td>53.8 (17.5)</td>
<td>55.0 (22.8)</td>
<td>48.0 (13.7)</td>
<td>54.2 (12.2)</td>
</tr>
<tr>
<td></td>
<td>TBF latency</td>
<td>726.4 (153.1)</td>
<td>645.2 (143.0)</td>
<td>674.4 (111.3)</td>
<td>703.4 (213.8)</td>
</tr>
<tr>
<td></td>
<td>Percent false alarms</td>
<td>35.7 (16.9)</td>
<td>40.2 (19.7)</td>
<td>38.1 (12.9)</td>
<td>33.2 (15.2)</td>
</tr>
<tr>
<td></td>
<td>Directed forgetting effect (%)</td>
<td>11.2 (15.2)</td>
<td>4.2 (25.1)</td>
<td>25.9 (12.3)</td>
<td>14.1 (17.7)</td>
</tr>
<tr>
<td></td>
<td>TBR percent correct</td>
<td>67.5 (21.8)</td>
<td>65.8 (18.4)</td>
<td>71.7 (14.7)</td>
<td>74.9 (13.9)</td>
</tr>
<tr>
<td></td>
<td>TBR latency</td>
<td>769.2 (188.5)</td>
<td>683.2 (151.8)</td>
<td>666.9 (96.3)</td>
<td>685.6 (171.5)</td>
</tr>
<tr>
<td></td>
<td>TBF percent correct</td>
<td>55.1 (23.9)</td>
<td>50.4 (22.6)</td>
<td>58.3 (14.2)</td>
<td>59.9 (16.9)</td>
</tr>
<tr>
<td></td>
<td>TBF latency</td>
<td>737.6 (149.7)</td>
<td>685.5 (119.3)</td>
<td>679.4 (95.4)</td>
<td>722.7 (186.9)</td>
</tr>
<tr>
<td></td>
<td>Percent false alarms</td>
<td>34.8 (20.6)</td>
<td>40.7 (20.6)</td>
<td>41.2 (12.4)</td>
<td>37.5 (17.5)</td>
</tr>
<tr>
<td></td>
<td>Directed forgetting effect (%)</td>
<td>12.5 (20.4)</td>
<td>15.4 (25.1)</td>
<td>13.3 (14.1)</td>
<td>14.9 (17.8)</td>
</tr>
<tr>
<td></td>
<td>TBR percent correct</td>
<td>64.9 (25.7)</td>
<td>65.2 (17.5)</td>
<td>71.3 (16.3)</td>
<td>75.1 (10.4)</td>
</tr>
<tr>
<td></td>
<td>TBR latency</td>
<td>699.9 (139.4)</td>
<td>667.2 (146.4)</td>
<td>620.6 (65.6)</td>
<td>674.9 (142.3)</td>
</tr>
<tr>
<td></td>
<td>TBF percent correct</td>
<td>42.3 (12.7)</td>
<td>44.5 (19.1)</td>
<td>46.0 (14.3)</td>
<td>50.7 (14.1)</td>
</tr>
<tr>
<td></td>
<td>TBF latency</td>
<td>749.5 (195.7)</td>
<td>646.9 (123.7)</td>
<td>697.7 (82.9)</td>
<td>716.3 (229.0)</td>
</tr>
<tr>
<td></td>
<td>Percent false alarms</td>
<td>18.0 (13.5)</td>
<td>24.5 (20.6)</td>
<td>24.6 (16.9)</td>
<td>18.5 (14.7)</td>
</tr>
<tr>
<td></td>
<td>Directed forgetting effect (%)</td>
<td>22.6 (20.1)</td>
<td>20.7 (15.4)</td>
<td>25.3 (18.5)</td>
<td>24.5 (17.8)</td>
</tr>
</tbody>
</table>

Pattern recognition memory

| Immediate                  | Percent correct                  | 92.8 (8.4)          | 90.6 (10.8) | 95.2 (7.3)          | 97.6 (5.6) |
| Latency                    | 1761.1 (511.3)                   | 1814.9 (506.7)     | 1576.5 (340.9) | 1532.2 (337.6) |

Delayed

| Percent correct | 84.4 (17.7) | 84.4 (13.5) | 92.9 (10.5) | 95.5 (6.8) |
| Latency         | 1759.3 (388.2) | 1710.6 (400.9) | 1712.5 (427.8) | 1706.9 (331.0) |

Affective Go/No-go

(words)

Happy shift

| Latency                | 500.6 (59.5) | 502.5 (90.4) | 486.0 (59.8) | 480.0 (61.5) |
| Commission errors     | 2.7 (3.2)   | 2.3 (2.3)   | 1.5 (1.2)   | 2.0 (1.8)   |
| Omission errors       | 0.87 (1.1)  | 1.2 (1.3)   | 0.47 (0.92) | 0.47 (0.74) |

Happy non-shift

| Latency                | 494.1 (60.7) | 496.9 (71.1) | 474.4 (62.3) | 482.0 (63.9) |
| Commission errors     | 2.2 (3.0)   | 2.5 (3.3)   | 1.6 (1.6)   | 1.3 (1.4)   |
| Omission errors       | 0.73 (1.2)  | 0.93 (1.5)  | 0.47 (0.83) | 0.33 (0.62) |

Sad shift

| Latency                | 508.6 (78.2) | 512.3 (98.9) | 486.8 (70.4) | 481.3 (55.7) |
| Commission errors     | 2.7 (2.6)   | 2.9 (2.5)   | 1.3 (1.3)   | 1.6 (1.8)   |
| Omission errors       | 0.40 (0.63) | 1.1 (2.0)   | 0.33 (0.62) | 0.13 (0.35) |

Sad non-shift

| Latency                | 501.6 (70.7) | 487.4 (77.8) | 489.9 (55.8) | 486.1 (78.4) |
| Commission errors     | 2.7 (3.2)   | 2.2 (3.1)   | 1.3 (1.0)   | 1.7 (1.4)   |
| Omission errors       | 0.33 (0.82) | 0.87 (1.2)  | 0.07 (0.26) | 0.40 (1.1)  |

TBR, To be remembered; TBF, to be forgotten.

Discussion

This study used a double-blind, placebo-controlled crossover design to investigate the effect of allelic
variation at the 5-HTTLPR on emotional and cognitive responses to ATD.

Cognitive measures

Memory and attention

The pattern of results at the recall stage of the ADF test was in line with the experimental predictions. The ss genotype group showed impaired recall performance following Trp−, while the ll genotype group showed no difference between Trp+ and Trp− conditions. To our surprise, following Trp+, the ss genotype group also showed a strong trend towards better recall performance than the ll genotype group. At the recognition stage of the ADF test, there was no significant difference in percent correct between the two genotype groups, nor was there any difference in the magnitude of the ‘directed forgetting effect’, and neither of these indices were affected by ATD.

The apparent episodic memory superiority in the ss genotype group was confirmed by the results of the
CANTAB PRM test. Again, the ss individuals were more accurate than ll individuals. The most pronounced effect of genotype on performance was evident at the delayed stage, although this could be due to a ceiling effect at the immediate stage. However, in contrast to expectations, no difference was apparent between the two genotype groups in terms of response to ATD on CANTAB PRM. A number of studies have examined memory performance following ATD, most of which have used recall paradigms. Two have reported impaired recall but intact recognition following ATD (Klaassen et al., 1999; Riedel et al., 1999), although others have reported impaired recognition performance following ATD (Rubinsztein et al., 2001; Schmitt et al., 2000). It is possible that recall is more sensitive than recognition to neurochemical manipulation, since recall places much higher demands on frontal retrieval mechanisms and is, therefore, more difficult (Squire, 1987).

On the AGNG test, the ss and ll genotype groups performed equivalently in terms of latency and commission errors, but the ss group made fewer omission errors than the ll group. None of these measures were affected by ATD. It is possible that improved vigilance in the ss group, reflected by fewer omission errors, might relate to the memory advantage observed at the recall stage of the Directed Forgetting test and the CANTAB PRM. In order to explore this hypothesis we performed post-hoc linear regression between omission errors on the AGNG test and directed forgetting recall and recognition and CANTAB PRM accuracy measures, averaging across treatment condition, including individuals of both genotypes. There was a significant negative relationship between omission error rate and accuracy at the immediate [F(1, 28) = 6.3, p = 0.02, adjusted \( r^2 = 0.16 \)] and delayed [F(1, 28) = 5.8, p = 0.02, adjusted \( r^2 = 0.15 \)] stages of the PRM, and a strong trend towards a significant negative relationship with accuracy at the recall stage of the ADF test [F(1, 27) = 5.4, p = 0.055, adjusted \( r^2 = 0.10 \)], although no relationship was apparent with accuracy at the recognition stage of the ADF test [F(1, 27) = 1.2, p = 0.3, adjusted \( r^2 = 0.008 \)]. While it is acknowledged that this post-hoc analysis is far from conclusive, it is consistent with the hypothesis that the s allele confers not only a more anxious temperament, a component of which is increased vigilance, but also, and as a result of this temperament, improved episodic memory. Improved mnemonic and attentional processes may provide carriers of the s allele with an evolutionary advantage, offsetting the clear evolutionary disadvantage of vulnerability to depressive illness and suicide in response to stressful events (Caspi et al., 2003).

**Emotional processing**

Based on previous experiments that suggested that ATD induced a bias towards negatively toned material, in the domains of both attention (Murphy et al., 2002) and memory (Klaassen et al., 2002), it had been predicted that the ss genotype group would display a negative bias on the ADF test and the AGNG tests. On the ADF test, it was predicted that ATD would cause ss participants difficulty in inhibiting encoding for sad words that they had been instructed to forget, producing a smaller directed forgetting effect for sad words than for happy or neutral words following ATD. On the AGNG tests, it was predicted that ss participants would show faster reactions to negative words than positive words following ATD. However, in neither case did ss participants show such a bias toward negative material following ATD.

Interestingly, on the Directed Forgetting test, regardless of genotype or treatment, participants showed a greater directed forgetting effect for neutral words than for emotionally toned words. This was attributable to better recognition of emotionally toned TBF words than neutral TBF words. One possible interpretation is that the presentation of emotionally toned words caused a change in emotional state in participants that interfered with the ability to inhibit encoding. Similar effects are also seen on Directed Forgetting tests in depressed patients, who show improved retrieval of negative TBF words (Power et al., 2000), and patients with obsessive–compulsive disorder (OCD), who show improved recognition of OCD-related TBF words (Tolin et al., 2002).

**Mood rating scales**

No differences on measures of internal state or depressive symptoms were noted in response to ATD between the ll and ss genotype groups. This is somewhat surprising, since the only comparable study to date reported that s-allele-carrying women were particularly vulnerable to mood change (also using the HDRS) following ATD (Neumeister et al., 2002). It has also previously been reported that ecstasy users of the ss genotype group are vulnerable to depressive symptoms following chronic use, which is thought to result in 5-HT depletion (Green et al., 2003; Roiser et al., 2005). The present study included both men and women, although there was not sufficient power to conduct a meaningful comparison between the genders. It is possible that the inclusion of male volunteers obscured effects in the ss genotype group, although this is unlikely as the mean HDRS score for ss females following ATD, was 2.8. Neumeister et al. (2002)
reported a mean HDRS score of ~9 in ss females following ATD, indicative of a clinically meaningful level of mood change.

**Study limitations and potential improvements**

Although a number of interesting findings were identified in the present study, some limitations must be acknowledged. First, in order to maximize potential differences between the genotypes, ls volunteers were not studied. This makes comparison with other studies employing ATD difficult. Since the most common genotype at the 5-HTTLPR in the Caucasian population is ls, it is likely that other studies that have examined the cognitive effects of ATD-recruited groups comprising predominantly of ls volunteers, and it is not known whether their behavioural response to ATD is markedly different from the ll and the ss genotype groups. Although Lesch et al. (1996) reported that ls heterozygotes resemble the ss genotype, it is also possible that the response of ls volunteers to ATD would be intermediate of those of the ll and ss genotypes, or even that ls volunteers may show a qualitatively different response to ATD. This question awaits experimental investigation.

Second, it was only possible to recruit 15 individuals of each genotype in this study, limiting the power of the study to detect differences of small effect size between the genotypes. In addition, since the s allele has been associated with reduced serotonergic synaptic transmission, the finding that ss volunteers outperformed ll volunteers on a number of indices is somewhat surprising. Given the small numbers of participants included in this study, the unexpected differences in cognitive performance between 5-HTTLPR genotypes should be treated with caution until replicated independently.

Third, participants of both positive and negative family history for depression were included in the present study, although the numbers in each genotype group were too small to make meaningful comparisons. Neumeister et al. (2002) reported mood change exclusively in healthy volunteers carrying the s allele, and that this effect interacted with family history of depression only in the ls genotype. Nevertheless, it would be of interest to determine whether a family history of depression and possession of the ss genotype interact to produce a particularly severe response to ATD.

Finally, recent studies of the 5-HTTLPR have described a new allele, termed Lg, that has transcriptional efficacy comparable to the s allele (Wendland et al., 2006). However, this allele had not been described when we began our pre-selection process, and therefore we did not genotype our participants for this polymorphism. Future studies of the effect of 5-HTTLPR polymorphism on response to pharmacological intervention should take this new allele into account in the analyses.

**Conclusion**

These results suggest a central role for 5-HT in episodic memory, and that ss individuals are particularly vulnerable to 5-HT depletion. However, further work is necessary to determine the precise mechanism by which the s allele confers vulnerability to 5-HT depletion. Polymorphism at the 5-HTTLPR appears to play an important role in determining individual response to ATD, and future studies employing ATD may benefit from stratifying results by genotype.

**Acknowledgements**

This work was supported by a Wellcome Trust programme grant to T. W. Robbins, B. J. Sahakian, B. J. Everitt and A. C. Roberts (No. 019407) and was completed within the University of Cambridge Behavioural and Clinical Neuroscience Institute, supported by a joint award of the Medical Research Council and Wellcome Trust. Many thanks to Caroline Humphries and all the staff at the Wellcome Trust Clinical Research Facility at Addenbrooke’s Hospital, Cambridge for their help and support. We thank Professor David C. Rubinsztein for supervising the genotyping and Dr Sonja Kotz for her help with stimulus selection for the ADF task. B. J. S. holds the F. C. Donders Chair in Psychopharmacology at Utrecht University. J. P. R. was supported by an MRC studentship and U. M. by a research fellowship of the Alexander von Humboldt Foundation. We also thank Dr M. Franklin and Dr J. Odontiadis (University of Oxford) for analysis of plasma data and Dr J. Whitaker for DNA extraction.

**Statement of Interest**

L. Clark, J. P. Roiser and B. J. Sahakian have performed consultancy work for Cambridge Cognition.

**References**


Park SB, Coull JT, McShane RH, Young AH, Sahakian BJ, Robbins TW, Cowen PJ (1994). Tryptophan depletion in normal volunteers produces selective...
impairments in learning and memory. *Neuropsychopharmacology* 33, 575–588.


