Improvement of executive functions in boys with attention deficit hyperactivity disorder: an open-label follow-up study with once-daily atomoxetine

Susan Shur-Fen Gau1,2 and Chi-Yung Shang1,2
1 Department of Psychiatry, National Taiwan University Hospital, Taipei, Taiwan
2 Department of Psychiatry, College of Medicine, National Taiwan University, Taipei, Taiwan

Abstract
Atomoxetine is efficacious in reducing symptoms of attention deficit hyperactivity disorder (ADHD) but its effect on executive functions needs more investigation. We examined the effect of atomoxetine on a wide range of non-verbal executive functions among 30 drug-naive male patients with DSM-IV ADHD, aged 8–16 yr, in an open-label 12-wk atomoxetine treatment trial. Before administration of atomoxetine, the participants were assessed by psychiatric interviews, the WISC-III, and the tasks involving executive functions of the Cambridge Neuropsychological Test Automated Battery (CANTAB): Intra-dimensional/Extra-dimensional Shifts (IED), Rapid Visual Information Processing (RVIP), Spatial Span (SSP), Spatial Working Memory (SWM), and Stockings of Cambridge (SOC); and reassessed at weeks 4 and 12. All the raw scores of the CANTAB were transformed to $z$ scores based on a normative sample of 180 children aged 8–16 yr. Results showed significant improvement in executive functions after treatment with atomoxetine for 4 wk or 12 wk including improved shifting and flexibility of attention in the IED; improved spatial short-term memory in the SSP; improved sustained attention and increased response inhibition in the RVIP; improved spatial working memory in the SWM; and improved spatial planning and problem solving in the SOC. Our findings suggested that atomoxetine was associated with significant improvement in various non-verbal executive functions among boys with ADHD, in addition to its well-known efficacy in ADHD-related symptom reductions. However, owing to lack of a placebo-controlled trial design, the findings should be interpreted with caution that changes in performance may be due to practice effects.

Introduction
Attention deficit hyperactivity disorder (ADHD) is a common childhood neuropsychiatric disorder (Gau et al. 2005; Polanczyk et al. 2007) with long-term academic and social impairments (Faraone et al. 2000), which may be mediated by impaired executive functioning (Diamantopoulou et al. 2007). The literature documents lifelong executive function deficits in ADHD (Seidman, 2006) with most consistent results for planning, working memory, and inhibition, followed by set-shifting tasks (Pennington & Ozonoff, 1996) with greater reductions in the visuo-spatial executive functions than the verbal working memory (Martinussen et al. 2005). Executive functions are regulated by catecholaminergic systems (Arnsten & Li, 2005) and are mediated by distinctive brain areas, including the prefrontal cortex, thalamus, basal ganglia, cerebellum, parietal and temporal lobes as...
Atomoxetine, a highly selective noradrenaline re-uptake inhibitor (SNRI), is a potent inhibitor of the presynaptic norepinephrine transporter, with little affinity for other noradrenergic receptors or for other neurotransmitter transporters (Simpson & Perry, 2003). Atomoxetine augments norepinephrine levels and indirectly increases dopamine levels in the pre-frontal cortex without increased catecholamine levels in nucleus accumbens (Bymaster et al. 2002), which may underlie the addictive properties of stimulants (Koob & Le Moal, 1997), the first approved medications by the Food and Drug Administration (FDA) in the USA for treating ADHD (MTA Cooperative Group, 1999). Atomoxetine was approved by the FDA as the first non-stimulant substance for treating ADHD in the USA in 2002. Previous studies have shown its efficacy in reducing clinical symptoms of ADHD (Caballero & Nahata, 2003; Gau et al. 2007b), lack of addictive potential (Bymaster et al. 2002), and potential clinical advantages in terms of sleep (Prasad & Steer, 2008) and comorbidities with tics (Allen et al. 2005) and anxiety (Geller et al. 2007).

Executive dysfunctions have been suggested as the independent targets for ADHD treatment (Faraone et al. 2005). While many studies support dopaminergic dysfunction underlying cognitive deficits in ADHD (Castellanos & Tannock, 2002), converging lines of data suggest that norepinephrine dysfunction may also contribute to cognitive impairments in ADHD (Newman et al. 2008; Viggiano et al. 2004). In contrast to the relatively well-known efficacy of atomoxetine in reducing ADHD symptoms (Caballero & Nahata, 2003; Gau et al. 2007b), little is known about its efficacy in executive dysfunction. Three animal studies showed that atomoxetine improved sustained spatial attention and response inhibition (five-choice serial reaction time test) (Blondeau & Delleu-Hagedorn, 2007), decreased impulsivity (stop-signal test) (Robinson et al. 2008), and remediated attentional deficits (set-shifting test) (Newman et al. 2008).

Human studies have demonstrated that a single clinically relevant oral dose of atomoxetine was associated with improved inhibitory control (stop-signal reaction time) in healthy participants (Chamberlain et al. 2006) and adults with ADHD (Chamberlain et al. 2007), and decreased commission errors (continuous performance test; CPT) in adults with ADHD (Chamberlain et al. 2007); and 3-wk and 10-wk treatments with atomoxetine increased inhibitory capacity (the Stroop Color Word Task) in adults with ADHD (Faraone et al. 2005; Spencer et al. 1998) but did not have effects on sustained attention (auditory CPT; Spencer et al. 1998), attentional set shifting (the Wisconsin Card Sorting Test; Spencer et al. 1998), visual memory (the Rey–Osterrieth Complex Figures Test; Spencer et al. 1998), or spatial working memory [the Spatial Working Memory (SWM) test of the Cambridge Neuropsychological Test Automated Battery (CANTAB)] (Chamberlain et al. 2007) in adults with ADHD. Despite the growing clinical use of atomoxetine in children, only one pilot study, published as a letter to the editor, reported a significant effect of atomoxetine on sustained attention in nine children and young people with ADHD (Barton et al. 2005). In summary, the reported effects of atomoxetine on the executive functions associated with ADHD are inconsistent and only one study was conducted in children, and this study had a relatively small sample size (n < 9) (Barton et al. 2005).

Hence, we conducted the present study to assess the long-term efficacy of atomoxetine on the improvements of executive functioning measured by five CANTAB tasks as the primary aim and on the symptom reductions as the secondary aim in 30 boys with ADHD. We used the CANTAB with good psychometric properties established in Western populations (Luciana, 2003; Luciana & Nelson, 1998) and in Taiwan (Gau et al. in pressa) because of its availability in parallel forms for repeated testing and its sensitivity to pharmacological intervention (Chamberlain et al. 2007; Coghill et al. 2007; Rhodes et al. 2005, 2006). We hypothesized that atomoxetine would improve performance on CANTAB tasks with prominent executive demands in addition to ADHD-related symptom reductions.

Method

Participants

We consecutively recruited 30 drug-naive male patients aged 8–15 yr (mean ± standard deviation, s.d., 10.70 ± 1.84) with clinically diagnosed DSM-IV ADHD from the Children’s Mental Health Center of National Taiwan University Hospital, Taipei. Most of their parents had college or higher educational levels (mothers 50%, fathers 60%) and were employed as technical personnel (mothers 50%, fathers 77%) followed by home-maker for the mothers (50%) and professional jobs for the fathers (23%).

The participants and their mothers were interviewed by using the Chinese version of the Schedule for Affective Disorders and Schizophrenia for School-Age Children – Epidemiological Version (K-SADS-E;
Gau et al. (2005) to confirm the participants’ DSM-IV diagnoses of ADHD and other psychiatric disorders. The Chinese K-SADS-E, a reliable and valid instrument and extensively used in a variety of studies regarding child and adolescent mental disorders in Taiwan (Gau & Soong, 1999; Gau et al. 2005, 2007a, b), was employed to make the DSM-IV diagnoses of ADHD and other psychiatric disorders. One of the authors (S.S.G.) made the best estimate of the diagnoses of ADHD and other psychiatric disorders based on K-SADS-E interviews of the participants and their mothers, and clinical assessments. The best-estimate procedure has been described in details elsewhere (Gau & Chiang, 2009).

Participants were excluded if they had a serious medical illness, such as a cardiovascular disease; had a full-scale IQ score <80; had a history of bipolar I or II disorders, psychosis, any substance abuse, or pervasive developmental disorder; had anxiety disorders based on DSM-IV criteria at study entry; had a history of any seizure disorder or prior electroencephalogram abnormalities related to epilepsy, or if they had ever used any psychoactive medications before the study.

Of the 30 patients with ADHD, 15 (50.0%) had combined type, 13 (43.3%) predominantly inattentive type, and two (6.5%) predominantly hyperactive-impulsive type. Eleven patients also met DSM-IV diagnostic criteria for other disorders, including oppositional defiant disorder (n = 11, 36.7%), conduct disorder (n = 1, 3.3%), and past history of anxiety disorders (n = 2, 6.7%).

Primary efficacy measures

Five CANTAB tasks involving executive abilities were employed in this study.

Intra-dimensional/Extra-dimensional Shifts (IED)

The IED assesses a subject’s ability to selectively maintain his/her attention on the specific attribute of compound stimuli across different examples, intra-dimensional shift, and then to shift their attention to a previously irrelevant attribute of stimuli, extra-dimensional shift (EDS) (Luciana & Nelson, 1998). The participants progressed through the tests by satisfying a set criterion of learning at each stage (six consecutive correct responses). If the participant failed to reach this criterion after 50 trials at any stage, the test was terminated. Five target indices were included: (1) pre-EDS errors: the number of errors made prior to the EDS stage (blocks 1–7); (2) EDS errors: errors made in the EDS stage (block 8); (3) completed stages: the number of stages successfully completed; (4) adjusted total errors: the adjusted score calculated by adding 25 for each stage not attempted due to failure (this value of 25 is used since the subjects must complete 50 trials to fail a stage and half of these could be correct by chance alone); and (5) adjusted total trials: the adjustment adds 50 for each stage not attempted due to failure at an earlier stage.

Rapid Visual Information Processing (RVIP)

The RVIP, a 4-min visual CPT modified and simplified from Wesnes & Warburton’s task (Wesnes & Warburton, 1984), is designed to assess sustained attention capacity (Sahakian et al. 1989). Digits (ranging from 2 to 9) appear one at a time (100 digits/min) in the centre of the screen in a random order. Participants were asked to press a response pad when they detected any one of three number sequences (3–5–7, 2–4–6, 4–6–8). The score of total hits represents the number of occasions upon which the target sequence was correctly responded to. The score of total misses represents the number of occasions the participant failed to respond to a target sequence within the response window. The score of total false alarms reports the number of times the participant responded outside the response window of a target sequence. The score of total correct rejections is the number of stimuli that were correctly rejected. Five indices were presented: (1) probability of hits (h, the participant responding correctly): total hits divided by the sum of total hits and total misses; (2) probability of false alarms (f, the participant responding inappropriately): total false alarms divided by the sum of total false alarms and total correct rejections; (3) A’ (is calculated as \[0.5 + \frac{[h - f] + (h - f)^2}{4 \times h \times (1 - f)}\]): a signal detection measure of sensitivity to the target, regardless of response tendency (Sahgal, 1987); (4) B’ (is calculated as \[\frac{(l - h^2) - (f - f^2)}{(l - h^2) + (f - f^2)}\]): a signal detection measure of the strength of trace required to elicit a response (Sahgal, 1987); and (5) mean latency: mean time taken to respond in correct responses.

Spatial Span (SSP)

The SSP measures spatial short-term memory. Similar to the Corsi blocks task (Milner, 1971), this task requires the ability to remember the order in which visual stimuli are presented. At the beginning, nine white boxes were presented in fixed locations on the screen. Next, the boxes changed colour, one after the other, in a pre-determined sequence, and the end of the sequence was indicated by a tone. Then, participants were asked to point to the boxes in the order in which they had changed colour. The test began with
2-box problems up to 9-box problems. Three indices were reported: (1) span length: the longest sequence successfully recalled; (2) total errors: the number of times an incorrect box was selected; and (3) total usage errors: the number of times a box selected was not in the sequence being recalled.

SWM

The SWM, based on a self-ordered search test (Petrides & Milner, 1982), an adaptation of Olton’s radial arm maze (Olton, 1987), assesses non-verbal working memory. Participants were asked to search through a number of coloured boxes presented on the screen for the blue tokens hidden inside. Only a single token was hidden in one of the boxes at each trial. Once found, there were no more tokens in this box within this problem. Two major indices were presented: (1) strategy utilization: the number of search sequences starting with a novel box in the difficult problems (both 6- and 8-box problems); (2) errors in the total and three different levels of difficulty (4-, 6-, and 8-box problems): the total errors for 4-, 6-, and 8-box problems were calculated based on the between-errors, within-errors, and double errors of particular box problems (i.e. between errors + within errors – double errors).

Stockings of Cambridge (SOC)

The SOC assesses spatial planning based on the Tower of London test (Shallice, 1982), and requires participants to move balls according to a goal position with given orders and locations. At the beginning, three suspended vertical stockings and three colored balls are presented on the monitor screen. Participants were required to move the colored balls, in a single move at a time, between the stockings to fulfil a goal position within a specified number of moves in the problem-solving condition; and subsequently, they were required to copy each move by following the identical sequence of moves played back by the computer, based on their employment of problem-solving in the control condition. The SOC comprises four problem sets (2-, 3-, 4-, and 5-move) reflecting increasing demands on planning. The active planning condition is the primary outcome measure, but the condition where subjects are asked to copy the moves on the screen is a motor control condition to adjust for motor timing which does not require higher level planning. Four major indices were presented: (1) problems solved in the specified minimum number of moves; (2) mean moves: the number of moves taken in excess of the specified minimum number, but within the maximum allowed; (3) initial thinking time: the difference in reaction time taken to select the first ball for the same problem under the two conditions; and (4) subsequent thinking time: the difference in time between selecting the first ball and completing the problem for the same problem under the two conditions.

Secondary efficacy measures

Clinical Global Impression-ADHD Severity (CGI-ADHD-S) rating

The CGI-ADHD-S is a single-item rating of the clinician’s assessment of the global severity of ADHD symptoms in relation to the clinician’s total experience with other ADHD patients. Severity was rated on a 7-point scale (from 1=normal, not at all ill, to 7=among the most extremely ill). The Chinese CGI-ADHD-S has been widely used in treatment studies on ADHD in Taiwan (Gau et al. 2007b, 2008a).

Chinese version of the Swanson, Nolan, and Pelham, version IV scale (SNAP-IV) – Parent Form

The SNAP-IV, a 26-item scale, consists of Inattention (items 1–9), Hyperactivity/Impulsivity (items 10–18), and Oppositionality (items 19–26), corresponding to the core symptoms of DSM-IV ADHD and oppositional defiant disorder, respectively (Swanson et al. 2001). The 26 items of SNAP-IV are rated on a 4-point Likert scale, with scores of 0–4 representing: ‘not at all’, ‘just a little’, ‘quite a bit’, and ‘very much’. The norms and psychometric properties of the Chinese version of the SNAP-IV (SNAP-IV-C) for parent (Gau et al. 2008b) and teacher (Gau et al. 2009) reports have been established in Taiwan.

Chinese Version of the Conners’ Parent Rating Scales – Revised:Short Form (CPRS-R:S)

The CPRS-R:S, a 27-item parent-reported rating scale, consists of three factor-derived subscales (those with the highest loadings on the CPRS-R:long form) and the ADHD index (Conners et al. 1998). The three subscales are Inattention/Cognitive Problems (six items), Hyperactivity/Impulsivity (six items), and Oppositional (six items). The ADHD index (12 items) is used to assess children and adolescents at risk for ADHD based on DSM-IV diagnostic criteria, rather than factor analysis. Each item is rated on a 4-point Likert scale (0 for never/seldom, 1 for occasionally, 2 for often/quite a bit, and 3 for very often/very frequently). The
Chinese version of the CPRS-R:S has been found to be a reliable and valid instrument for measuring ADHD-related symptoms in Taiwan (Gau et al. 2006); z scores were computed using the norm established in Taiwan (Gau et al. 2006).

**Procedures**

The Research Ethics Committee of National Taiwan University Hospital approved this study prior to the administration of any study procedure or the dispensing of study drugs. The procedures and purpose of this study, as well as reassurance of confidentiality, were clearly explained to the participants and their parents. Written informed consent was obtained from the participants and their parents. The study was monitored by the Research Ethics Committee at all stages, from inception to completion, in accordance with current good clinical practice in Taiwan.

The participants were assessed by the Weschler Intelligence Scale for Children – 3rd edition (WISC-III) to confirm their IQ > 80 and by the Chinese K-SADS-E for the diagnosis of ADHD and other psychiatric disorders prior to recruitment. Digit span was included in the WISC-III to assess sustained attention (digit forward) and verbal working memory (digit backward). At baseline (week 0), the participants were assessed by the CANTAB, and clinically evaluated by the authors using the CGI-ADHD-S. The participants' mothers reported their ADHD-related symptoms on the self-administered measures, the Chinese CPRS-R:S and SNAP-IV. The participants were reassessed with the CANTAB, the investigator-administered CGI-ADHD-S, and maternal reports on the Chinese CPRS-R:S and SNAP-IV on week 4 (visit 2) ± 3 days and week 12 (visit 3) ± 3 days after treatment with atomoxetine. The initial once-daily dose of atomoxetine was 0.5 mg/kg lasting for 7 days, followed by increasing the dose of atomoxetine to 1.2 mg/kg on day 8 without further changes in doses throughout the study period to endpoint. Safety was assessed by the administration of open-ended questions for adverse events and by regular monitoring of vital signs and body weight at each of the three visits (weeks 0, 4, 12).

**Statistical analyses**

SAS version 9.1 (SAS Institute Inc., USA) was used to conduct data analysis. Alpha value was pre-selected at the level of p < 0.05. Using a sample of 180 normally developing children, grades 3–10 (ages 8–16 yr; mean ± s.d., 12.42 ± 1.62 yr), without lifetime DSM-IV diagnosis of ADHD based on the Chinese K-SADS-E as the norm for the CANTAB assessments, we transformed the raw score of each parameter to its z score, which has a mean of zero and a s.d. of 1, by using the mean and s.d. derived from the three age groups 8–10, 11–13, and 14–16 yr of this normative sample. This normative sample (134 boys, 46 girls) was derived from the school comparison group for probands with ADHD in a family study of ADHD (Gau & Chiang, 2009; Gau et al. in pressb). The data of the CANTAB were expressed by mean ± s.d. of the raw scores and z scores; the data of the CGI-ADHD-S and self-administered measures (SNAP-IV and CPRS-R:S) were expressed by mean ± s.d. of the raw scores and z scores, respectively.

Because of repeated measures within the same subject, we used a linear multi-level model to test the mean differences in the repeated measures of the executive functions measured by the CANTAB tasks, the CPRS-R:S, the SNAP-IV and the CGI-ADHD-S at weeks 4 and 12, compared to baseline (week 0), and to test the interactions between visits and the task difficulties. In addition, Cohen's d was used to compute the effect size (standardized difference between the two means) for the comparisons among the three assessments (baseline, week 4, week 12) with small, medium, and large effect sizes as Cohen's d ≥ 0.2 to < 0.5, ≥ 0.5 to < 0.8, and ≥ 0.8, respectively.

**Results**

**Sample characteristics, body weight, and vital signs**

Table 1 summarizes the sample characteristics and the changes of vital signs. The average full-scale IQ ranged from 80 to 135 (105.37 ± 13.55). After the initial atomoxetine doses for 7 days, all the participants, subsequently, received a mean daily dose of 1.20 ± 0.07 mg/kg without further changes of daily dose (p = 0.756, Table 1). Results comparing endpoint (week 12) to baseline showed no significant changes in body weight (p = 0.309), in systolic blood pressure (p = 0.523), in diastolic blood pressure (p = 0.820) and in pulse rate (p = 0.097, Table 1).

**Improvement of non-verbal executive function**

Table 2(a–d) summarizes the comparisons of raw scores and z scores, mean z score differences (95% CI), and the effect sizes (Cohen’s d) of the five non-verbal executive tasks of the CANTAB across the three visits (baseline, week 4, week 12).

IED. The pre-EDS errors, EDS errors, and adjusted total trials and errors significantly decreased at week 4 from baseline with medium effect sizes. The EDS
errors and adjusted total errors were significantly lower at week 12 than at baseline with small- to-medium effect sizes (Table 2a). These results suggested that atomoxetine was associated with improving set-shifting and flexibility of attention vs. baseline.

**RVIP.** Children with ADHD had higher probability of hits, better sensitivity to the target sequences, and shorter mean latency to respond correctly at weeks 4 and 12 than at baseline (absolute Cohen’s $d$ ranging from 0.38 to 0.80); they had lower probability of false alarm and needed stronger trace to elicit a response (Table 2b). They had lower probability of false alarm [Cohen’s $d = -0.39$, mean $z$ score difference $-0.78$ (95% CI $-1.34$ to $-0.19$)], and shorter mean latency to respond correctly [Cohen’s $d = -0.39$, mean $z$ score difference $-0.51$ (95% CI $-0.79$ to $-0.23$)] at week 12 than at week 4. These results indicate that atomoxetine was associated with improvement in sustained attention and response inhibition vs. baseline.

**SSP.** Children with ADHD had longer span sequences successfully recalled at week 12, and fewer total usage errors at weeks 4 and 12 than at baseline with medium effect sizes (Table 2c), suggesting improved spatial short-term memory with atomoxetine mainly after 12 wk of treatment.

**SWM.** Children with ADHD had better strategy use (lower strategy scores) to improve searching efficacy and had fewer total errors at week 12 than at baseline with small effect sizes (Table 2c). There was marginally significant interaction between the three visits and the task difficulties (4-, 6-, and 8-box problems) on the total errors ($F_{1,29} = 3.13$, $p = 0.097$).

**SOC.** The participants had fewer moves, more problems solved in the minimum number of moves, and shorter initial and subsequent thinking time at week 12 than at baseline (Cohen’s $d$ ranging from $0.35$ to $1.12$) (Table 2d). Moreover, they had more problems solved in the minimum number of moves at week 12 than at week 4 [Cohen’s $d = 0.57$, mean $z$ score difference $0.58$ (95% CI $0.26$–$0.90$)]. Their initial and subsequent thinking time was shorter at week 4 compared to baseline. There were marginally significant interaction between the three visits and the task difficulties (2-, 3-, 4-, and 5-move problems) on the mean moves ($b = 0.10$, $p = 0.053$) and the subsequent thinking time ($b = 2.19$, $p = 0.044$). The results indicate that atomoxetine was associated with improvement in spatial planning and problem solving after 12 wk treatment, and magnitude of improvement might increase with increased task difficulties.

**ADHD-related symptom reductions**

At baseline, the mean CGI-ADHD-S rating was ‘markedly ill’ to ‘severely ill’ (5.57 ± 0.73). The score significantly decreased to 3.43 (‘mildly ill’ to ‘moderately ill’) at week 4, and to 2.83 (‘borderline ill’ to
### Table 2a. Comparisons of performances in the Intra-dimension/Extra-dimension Shift test across baseline, week 4 and week 12

<table>
<thead>
<tr>
<th>Raw score (z score)</th>
<th>Baseline (mean ± S.D.)</th>
<th>Week 4 (mean ± S.D.)</th>
<th>Week 12 (mean ± S.D.)</th>
<th>Week 4 (baseline)</th>
<th>Week 12 (baseline)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>β (95% CI)</td>
<td>β (95% CI)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>p value</td>
<td>p value</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>d</td>
<td>d</td>
</tr>
<tr>
<td>Extra-dimensional</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>shift errors</td>
<td>15.5 ± 11.3 (0.74 ± 1.23)</td>
<td>10.2 ± 10.6 (0.16 ± 1.15)</td>
<td>9.3 ± 10.1 (0.07 ± 1.09)</td>
<td>-0.58 (-1.11 to -0.05)</td>
<td>-0.49 (-0.34 to -0.56)</td>
</tr>
<tr>
<td>Pre-extra-dimensional</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>shift errors</td>
<td>9.8 ± 6.1 (0.70 ± 1.52)</td>
<td>7.1 ± 3.0 (0.03 ± 0.75)</td>
<td>8.9 ± 5.5 (0.47 ± 1.37)</td>
<td>-0.68 (-1.21 to -0.15)</td>
<td>-0.56 (-0.12 to 0.18)</td>
</tr>
<tr>
<td>Completed stages</td>
<td>7.8 ± 1.4 (-0.98 ± 1.68)</td>
<td>8.3 ± 0.9 (-0.36 ± 1.07)</td>
<td>8.4 ± 1.4 (-0.32 ± 1.69)</td>
<td>0.62 (-0.04 to 1.28)</td>
<td>0.44 (0.33 to 0.71)</td>
</tr>
<tr>
<td>Total errors (adjusted)</td>
<td>44.6 ± 34.0 (1.02 ± 1.57)</td>
<td>28.3 ± 22.0 (0.26 ± 1.01)</td>
<td>31.1 ± 34.9 (0.39 ± 1.61)</td>
<td>-0.75 (-1.36 to -0.14)</td>
<td>-0.63 (-0.31 to -0.01)</td>
</tr>
<tr>
<td>Total trials (adjusted)</td>
<td>133.5 ± 58.6 (1.07 ± 1.52)</td>
<td>104.4 ± 38.8 (0.32 ± 1.00)</td>
<td>111.4 ± 63.4 (0.50 ± 1.64)</td>
<td>-0.75 (-1.34 to -0.16)</td>
<td>-0.57 (-0.29 to 0.02)</td>
</tr>
</tbody>
</table>

S.D., Standard deviation; β, estimate of mean difference of z score; CI, confidence interval; d, Cohen’s d.

### Table 2b. Comparisons of performances in the Rapid Visual Information Processing test across baseline, week 4 and week 12

<table>
<thead>
<tr>
<th>Raw score (z score)</th>
<th>Baseline (mean ± S.D.)</th>
<th>Week 4 (mean ± S.D.)</th>
<th>Week 12 (mean ± S.D.)</th>
<th>Week 4 (baseline)</th>
<th>Week 12 (baseline)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>β (95% CI)</td>
<td>β (95% CI)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>p value</td>
<td>p value</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>d</td>
<td>d</td>
</tr>
<tr>
<td>Probability of hits</td>
<td>0.40 ± 0.22 (-1.05 ± 1.20)</td>
<td>0.51 ± 0.21 (-0.46 ± 1.16)</td>
<td>0.50 ± 0.26 (-0.47 ± 1.43)</td>
<td>0.59 (0.21 to 0.97)</td>
<td>0.51 (0.29 to 0.50)</td>
</tr>
<tr>
<td>Probability of false alarms</td>
<td>0.03 ± 0.03 (1.19 ± 1.82)</td>
<td>0.03 ± 0.04 (1.27 ± 2.35)</td>
<td>0.02 ± 0.03 (0.50 ± 1.54)</td>
<td>0.07 (-0.81 to 0.96)</td>
<td>0.03 (0.43 to 0.66)</td>
</tr>
<tr>
<td>A'</td>
<td>0.82 ± 0.08 (-1.44 ± 1.59)</td>
<td>0.85 ± 0.08 (-0.78 ± 1.51)</td>
<td>0.86 ± 0.08 (-0.58 ± 1.54)</td>
<td>0.65 (0.27 to 1.04)</td>
<td>0.38 (0.43 to 0.99)</td>
</tr>
<tr>
<td>B'</td>
<td>0.75 ± 0.20 (-1.47 ± 1.69)</td>
<td>0.79 ± 0.21 (-1.17 ± 1.77)</td>
<td>0.85 ± 0.16 (-0.66 ± 1.37)</td>
<td>0.26 (-0.28 to 0.80)</td>
<td>0.20 (0.42 to 0.73)</td>
</tr>
<tr>
<td>Mean latency (ms)</td>
<td>639 ± 201 (1.54 ± 1.72)</td>
<td>562 ± 171 (0.87 ± 1.46)</td>
<td>502 ± 136 (0.36 ± 1.16)</td>
<td>-0.66 (-1.14 to -0.19)</td>
<td>-0.42 (-0.39 to -0.84)</td>
</tr>
</tbody>
</table>

S.D., Standard deviation; β, estimate of mean difference of z score; A', sensitivity to errors; B', strength of trace required to elicit a response; CI, confidence interval; d, Cohen’s d.
Table 2c. Comparisons of performances in the Spatial Span and Spatial Working Memory tests across baseline, week 4 and week 12

<table>
<thead>
<tr>
<th>Raw score (z score)</th>
<th>Baseline (mean ± S.D.)</th>
<th>Week 4 (mean ± S.D.)</th>
<th>Week 12 (mean ± S.D.)</th>
<th>Week 4 (baseline)</th>
<th>Week 12 (baseline)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>β (95% CI)</td>
<td>p value</td>
<td>d</td>
<td>p value</td>
</tr>
<tr>
<td><strong>Spatial span</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Span length</td>
<td>5.9 ± 1.4 (–1.14 ± 1.05)</td>
<td>6.4 ± 1.5 (–0.78 ± 1.09)</td>
<td>6.8 ± 1.4 (–0.46 ± 1.04)</td>
<td>0.36 (–0.07 to 0.79)</td>
<td>0.34 (–0.16 to 0.52)</td>
</tr>
<tr>
<td>Total errors</td>
<td>13.4 ± 6.3 (0.07 ± 0.98)</td>
<td>13.8 ± 6.9 (0.14 ± 1.03)</td>
<td>13.2 ± 6.7 (0.05 ± 1.01)</td>
<td>0.07 (–0.39 to 0.53)</td>
<td>0.07 (–0.25 to 0.23)</td>
</tr>
<tr>
<td>Total usage errors</td>
<td>2.5 ± 1.7 (0.77 ± 1.26)</td>
<td>1.9 ± 1.3 (0.31 ± 0.98)</td>
<td>1.4 ± 1.5 (–0.05 ± 1.10)</td>
<td>–0.46 (–0.81 to –0.11)</td>
<td>–0.40 (–0.65 to –0.17)</td>
</tr>
<tr>
<td><strong>Spatial working memory</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total errors</td>
<td>34.1 ± 20.0 (1.05 ± 1.36)</td>
<td>31.9 ± 21.5 (0.91 ± 1.46)</td>
<td>27.5 ± 19.3 (0.61 ± 1.31)</td>
<td>–0.15 (–0.61 to 0.32)</td>
<td>–0.10 (–0.42 to –0.03)</td>
</tr>
<tr>
<td>Strategy utilization</td>
<td>34.5 ± 6.6 (0.48 ± 1.38)</td>
<td>32.7 ± 5.9 (0.11 ± 1.24)</td>
<td>31.7 ± 5.4 (–0.11 ± 1.12)</td>
<td>–0.38 (–0.83 to 0.08)</td>
<td>–0.29 (–0.48 to –0.11)</td>
</tr>
</tbody>
</table>

s.d., Standard deviation; β, estimate of mean difference of z score; CI, confidence interval; d, Cohen’s d.

Table 2d. Comparisons of performances in the Stockings of Cambridge test across baseline, week 4 and week 12

<table>
<thead>
<tr>
<th>Raw score (z score)</th>
<th>Baseline (mean ± S.D.)</th>
<th>Week 4 (mean ± S.D.)</th>
<th>Week 12 (mean ± S.D.)</th>
<th>Week 4 (baseline)</th>
<th>Week 12 (baseline)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>β (95% CI)</td>
<td>p value</td>
<td>d</td>
<td>p value</td>
</tr>
<tr>
<td>Problems solved in minimum moves</td>
<td>7.6 ± 2.0 (–0.54 ± 1.01)</td>
<td>7.7 ± 2.2 (–0.51 ± 1.11)</td>
<td>8.8 ± 1.8 (0.07 ± 0.91)</td>
<td>0.03 (–0.47 to 0.54)</td>
<td>0.03 (0.10 to 0.51)</td>
</tr>
<tr>
<td>Mean moves</td>
<td>18.2 ± 2.1 (0.56 ± 0.97)</td>
<td>17.6 ± 2.2 (0.26 ± 1.00)</td>
<td>17.0 ± 2.0 (–0.01 ± 0.91)</td>
<td>–0.29 (–0.78 to 0.19)</td>
<td>–0.30 (–0.50 to –0.06)</td>
</tr>
<tr>
<td>Mean initial thinking time (ms)</td>
<td>3981 ± 2894 (–0.32 ± 1.03)</td>
<td>3128 ± 1981 (–0.62 ± 0.71)</td>
<td>3069 ± 2309 (–0.65 ± 0.82)</td>
<td>–0.30 (–0.63 to 0.03)</td>
<td>–0.34 (–0.31 to –0.01)</td>
</tr>
<tr>
<td>Mean subsequent thinking time (ms)</td>
<td>1682 ± 1510 (0.77 ± 1.31)</td>
<td>753 ± 945 (–0.04 ± 0.82)</td>
<td>452 ± 370 (–0.30 ± 0.32)</td>
<td>–0.81 (–1.32 to –0.29)</td>
<td>–0.74 (–0.77 to –0.30)</td>
</tr>
</tbody>
</table>

s.d., Standard deviation; β, estimate of mean difference of z score; CI, confidence interval; d, Cohen’s d.
‘mildly ill’) at week 12 (Table 3), with a significant linear trend of decreasing symptom severity.

Regarding parental ratings, children with ADHD had significant score reductions in the Chinese CPRS-RS and SNAP-IV – Parent forms from baseline to week 4 (Cohen’s $d = -0.51$ to $-0.90$) and week 12 (Cohen’s $d = 0.80$ to $-1.15$) except in the oppositional subscales, which showed significantly reductions at week 12 with small effect sizes (Table 3).

**Discussion**

The present study is the first to examine the efficacy of atomoxetine in children with ADHD using a wide-range of executive tasks of the CANTAB with a greater sample size, longer follow-up, and more visits than previous human studies (Barton et al. 2005; Chamberlain et al. 2006, 2007; Faraone et al. 2005; Spencer et al. 1998). The major finding was that atomoxetine was associated with improvements vs. baseline in a variety of non-verbal executive functions in boys with ADHD including sustained attention (RVIP), inhibitory ability (RVIP), and attentional set shifting (IED) noted at week 4; and spatial short-term memory (SSP), spatial working memory (SWM), spatial planning (SOC) and spatial problem solving (SOC), mainly noted at week 12, although decreased usage errors in the SSP and reduced subsequent thinking time in the SOC were also noted at week 4. Moreover, the magnitude of improvement in spatial planning and problem solving may be a function of treatment duration of atomoxetine and task difficulties. Hence, our results lend strong evidence to support the findings from animal studies (Blondeau & Dellu-Hagedorn, 2007; Newman et al. 2008; Robinson et al. 2008), and previous human studies (Barton et al. 2005; Chamberlain et al. 2006, 2007; Faraone et al. 2005; Spencer et al. 1998) that atomoxetine may be an effective treatment for the executive dysfunction associated with ADHD, not only in Western populations (Barton et al. 2005; Chamberlain et al. 2007), but also in an ethnic Chinese population, and not only in adults (Chamberlain et al. 2006, 2007) but also in children. Like others (Faraone et al. 2005; Spencer et al. 1998), we did not find that chronic administration of atomoxetine adversely affected executive functions in ADHD children. Similar to previous clinical trials (Bangs et al. 2008; Caballero & Nahata, 2003; Gau et al. 2007b), our findings demonstrate a reduction in ADHD-related symptoms across the three visits based on two well-validated parental rating scales (Gau et al. 2006, 2008b) and investigator’s assessments; however, lack of a control group has prevented us from making any conclusions about the effectiveness of atomoxetine in reducing ADHD-related symptoms in the present study.

Although Spencer & colleagues (1998) reported that atomoxetine did not improve auditory CPT, our results showed that atomoxetine significantly increased sustained attention measured by visual CPT. This finding is consistent with the norepinephrine hypothesis on regulation of attention (Coull et al. 2004; De Martino et al. 2008) and lends evidence to support the finding from a pilot study (Barton et al. 2005). Moreover, consistent with others (Barton et al. 2005; Chamberlain et al. 2006, 2007), atomoxetine was associated with decreased probability of false alarms and increased strength of trace required to elicit a response, consistent with its efficacy in reducing impulsivity (Robinson et al. 2008) and improving inhibitory controls (Barton et al. 2005; Spencer et al. 1998).

Unlike the negative findings in adults with ADHD (Chamberlain et al. 2007; Spencer et al. 1998), the present study demonstrated that the set-shifting performance in children with ADHD deviated from that of the normative sample, and the beneficial effects of atomoxetine on cognitive flexibility measured by the IED. Evidence from animal studies has suggested that the noradrenergic system modulates cognitive flexibility (Lapiz & Morilak, 2006; Newman et al. 2008; Tait et al. 2007) by showing that decreased and increased norepinephrine in the prefrontal cortex impaired (Lapiz & Morilak, 2006; Tait et al. 2007) and improved (Newman et al. 2008) attentional set shifting, respectively.

Despite no effect of atomoxetine in improving visual memory (Spencer et al. 1998), and spatial working memory (Chamberlain et al. 2007) in adults with ADHD, the present study found an improvement in spatial short-term memory and working memory across the three visits in boys with ADHD treated with atomoxetine. Animal studies have shown that norepinephrine efflux in the prefrontal cortex is selectively increased during a task measuring spatial working memory (Rossetti & Carboni, 2005), suggesting that norepinephrine may be involved in the active maintenance of spatial information (Rossetti & Carboni, 2005) and influence working memory processes (Arnstien & Li, 2005). Another novel finding of our study is the improvement in spatial planning and problem solving measured by the CANTAB, with greater improvement magnitude with increasing task difficulties and treatment duration. This finding was partially supported by a healthy adult study revealing the noradrenergic modulation in problem solving on difficult tasks (Campbell et al. 2008). The effect of
Table 3. Comparisons of ADHD-related symptoms across baseline, week 4 and week 12

<table>
<thead>
<tr>
<th></th>
<th>Baseline (mean ± S.D.)</th>
<th>Week 4 (mean ± S.D.)</th>
<th>Week 12 (mean ± S.D.)</th>
<th>Week 4 (baseline difference)</th>
<th>Week 12 (baseline difference)</th>
<th>Linear</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Mean ± (95% CI)</td>
<td>p value</td>
<td>d</td>
</tr>
<tr>
<td>CGI-ADHD-S (rating 1–7)</td>
<td>5.57 ± 0.73</td>
<td>3.43 ± 0.90</td>
<td>2.83 ± 0.87</td>
<td>-2.13 (-2.53, -1.74)</td>
<td>-2.61</td>
<td>-1.37 (-1.54 to -1.19)</td>
</tr>
<tr>
<td>SNAP-IV – Parent Report, raw score (z scores)</td>
<td></td>
<td></td>
<td></td>
<td>p &lt; 0.001</td>
<td></td>
<td>p &lt; 0.001</td>
</tr>
<tr>
<td>Inattentive</td>
<td>17.03 ± 5.54</td>
<td>12.60 ± 4.20</td>
<td>10.93 ± 5.11</td>
<td>-1.16 (-1.66, -0.66)</td>
<td>-0.90</td>
<td>-0.82 (-1.10 to -0.53)</td>
</tr>
<tr>
<td></td>
<td>(2.93 ± 1.45)</td>
<td>(1.77 ± 1.10)</td>
<td>(1.33 ± 1.34)</td>
<td>p &lt; 0.001</td>
<td></td>
<td>p &lt; 0.001</td>
</tr>
<tr>
<td>Hyperactive-impulsive</td>
<td>11.72 ± 7.05</td>
<td>8.46 ± 5.57</td>
<td>7.04 ± 4.27</td>
<td>-1.03 (-1.51, -0.55)</td>
<td>-0.51</td>
<td>-0.72 (-1.03 to -0.42)</td>
</tr>
<tr>
<td></td>
<td>(2.81 ± 2.23)</td>
<td>(1.77 ± 1.76)</td>
<td>(1.32 ± 1.35)</td>
<td>p &lt; 0.001</td>
<td></td>
<td>p &lt; 0.001</td>
</tr>
<tr>
<td>Oppositional</td>
<td>10.34 ± 4.89</td>
<td>9.20 ± 4.15</td>
<td>8.14 ± 4.87</td>
<td>-0.28 (-0.68, 0.12)</td>
<td>-0.25</td>
<td>-0.28 (-0.50 to -0.06)</td>
</tr>
<tr>
<td></td>
<td>(1.44 ± 1.23)</td>
<td>(1.15 ± 1.04)</td>
<td>(0.89 ± 1.22)</td>
<td>p = 0.176</td>
<td></td>
<td>p = 0.020</td>
</tr>
<tr>
<td>CPRS-RS, raw score (z scores)</td>
<td></td>
<td></td>
<td></td>
<td>p = 0.002</td>
<td></td>
<td>p = 0.001</td>
</tr>
<tr>
<td>Inattentive/cognitive problems</td>
<td>10.53 ± 3.79</td>
<td>8.07 ± 3.95</td>
<td>6.93 ± 3.92</td>
<td>-0.84 (-1.34, -0.35)</td>
<td>-0.63</td>
<td>-0.62 (-0.87 to -0.37)</td>
</tr>
<tr>
<td></td>
<td>(2.73 ± 1.30)</td>
<td>(1.89 ± 1.36)</td>
<td>(1.50 ± 1.35)</td>
<td>p = 0.002</td>
<td></td>
<td>p = 0.001</td>
</tr>
<tr>
<td>Hyperactivity/impulsivity</td>
<td>8.30 ± 5.02</td>
<td>5.26 ± 3.44</td>
<td>4.77 ± 3.40</td>
<td>-1.45 (-2.02, -0.88)</td>
<td>-0.71</td>
<td>-0.84 (-1.15 to -0.54)</td>
</tr>
<tr>
<td></td>
<td>(3.35 ± 2.39)</td>
<td>(1.90 ± 1.64)</td>
<td>(1.67 ± 1.62)</td>
<td>p &lt; 0.001</td>
<td></td>
<td>p &lt; 0.001</td>
</tr>
<tr>
<td>Oppositional</td>
<td>6.73 ± 3.12</td>
<td>5.87 ± 2.92</td>
<td>5.57 ± 3.15</td>
<td>-0.38 (-0.92, 0.17)</td>
<td>-0.29</td>
<td>-0.25 (-0.48 to -0.02)</td>
</tr>
<tr>
<td></td>
<td>(1.95 ± 1.35)</td>
<td>(1.57 ± 1.26)</td>
<td>(1.44 ± 1.36)</td>
<td>p = 0.189</td>
<td></td>
<td>p = 0.041</td>
</tr>
</tbody>
</table>

S.D., standard deviation; CI, confidence interval; d, Cohen’s d; β, estimate of regression coefficient for linear trend; CGI-ADHD-S, Clinical Global Impressions-ADHD-Severity; SNAP-IV, Chinese Version of the Swanson, Nolan, and Pelham IV scale; CPRS-RS, Conners’ Parent Rating Scales – Revised; Short Form.

*Mean difference of z score.
atomoxetine in spatial short-term memory, working memory, planning, and problem-solving did not significantly emerge until week 12, rather than week 4 as shown in the sustained attention, inhibitory capacity, and set-shifting tasks.

No significant change of body weight and vital signs from baseline to follow-up as demonstrated in this study and in a randomized, double-blind, placebo-controlled study in Taiwan (Gau et al. 2007b) suggests that atomoxetine may be safe and well tolerated in Taiwanese children. Although weight loss is the most frequent finding in the physical measures of patients treated with atomoxetine (Michelson et al. 2001; Weiss et al. 2005), long-term studies have demonstrated that weight loss, an acute effect, does not persist during periods of extended treatment (Gau et al. 2007b; Spencer et al. 2005). Despite our negative findings in heart rate and blood pressure, other reports of modest increases in noradrenergic tone in patients treated with atomoxetine implies the importance of regular monitoring of vital signs among patients treated with atomoxetine (Chamberlain et al. 2007; Kelsey et al. 2004).

**Methodological considerations**

The strengths of the present study include being the first to examine the effect of atomoxetine on executive functions in children with ADHD with a larger sample size, longer follow-up period and more frequent assessments than other human studies; clinical and standardized psychiatric assessments based on child and mother interviews to inform psychiatric diagnosis; and comprehensive assessments of a wide range of executive functions using standardized and well-validated neuropsychological tests with a well-established group of children to transform the raw scores to $z$ scores (Gau et al. in pressb; Luciana, 2003; Luciana & Nelson, 1998).

The major limitation of our study was lack of placebo-controlled trial design, making the findings of improvement on executive functions with atomoxetine treatment much less convincing. As a repeated measure design, the vulnerability of executive functioning tests to factors as loss of novelty and learning effects has been highlighted (Lowe & Rabbitt, 1998). Although high temporal stability of the five tasks in the CANTAB (ICC ranging from 0.42 to 1.00 for 1-month test–retest reliability) has been demonstrated in our previous study, the changes in performance across the three repeated CANTAB assessments may not be solely attributable to the effect of atomoxetine but also due to learning effects. The learning effects are particularly problematic in the IED test because the extra-dimensional shift is likely to be recalled by subjects at future testing sessions leading to fewer errors at the extra-dimensional shift stage. Hence, we need to interpret the findings with caution due to the possible learning effects from repeated CANTAB assessments.

Next, although the sample size of this study is the largest among studies of atomoxetine efficacy on executive functions regardless of children or adults, the small sample size and male subjects only have limited our ability to examine the differential efficacy of atomoxetine on executive function as a function of sex, ADHD subtypes or comorbid patterns. Furthermore, a longer follow-up period of up to 12 wk compared to previous studies minimized the practice effects. Last, although this study demonstrated that atomoxetine may be associated with improvement in executive functions, a placebo-controlled, blinded clinical trial and a head-to-head comparison study of atomoxetine with psychostimulants warrants further investigation and functional brain-imaging studies are needed to explore the precise effects of atomoxetine on the neural circuitry of executive function (Chamberlain et al. 2009).

**Implications**

Our findings indicate that in addition to symptom reductions, atomoxetine is also associated with improvements in a variety of non-verbal executive functions with significant improvement noted after 4 wk of treatment for sustained attention, inhibitory control and set-shifting, and after 12 wk for spatial short-term memory, working memory, planning and problem-solving. However, their performance at endpoint is still not normalized. Therefore, long-term administration of atomoxetine is recommended for improving executive functions, which may mediate the amelioration of the academic performance and social functioning in children with ADHD, particularly when they face difficult and complicated tasks and situations. Regarding the research implication, a double-blind, placebo-controlled trial, to follow-up these data, would be valuable.

**Acknowledgements**

This work was supported by grants from the National Health Research Institute (NHRI-EX97-9407PC), Taiwan, and an investigator-initiated study sponsored by Eli Lilly & Co. Taiwan (B4Z-TW-O021), which provided study drug (atomoxetine). The authors express thanks to Ming-Fang Chen, M.S., for assistance in data analysis; and to the anonymous
reviewers' helpful comments to strengthen the ultimate result of this paper. (ClinicalTrials.gov number: NCT00529893.)

Statement of Interest

Dr S. S. Gau and Dr C. Y. Shang have conducted clinical trials on behalf of Janssen-Cilag and Eli Lilly & Co., Taiwan; Dr S. S. Gau was on the speakers’ bureau and a consultant for Janssen-Cilag and Eli Lilly & Co., Taiwan; and was on the speakers’ bureau for Astellas Pharma Inc., Taiwan.

References


Gau SS, Chong MY, Yang P, Yen CF, et al. (2007a). Psychiatric and psychosocial predictors of substance use...


