Original Paper

Combined neuropsychological and neurophysiological assessment of drug effects on groups and individuals

Alan Gevins, Aaron B Ilan, An Jiang, Lita Sam-Vargas, Cliff Baum and Cynthia S Chan

Psychopharm

Journal of Psychopharmacology
25(8) 1062–1075

DOI: 10.1177/0269881110388334
jop.sagepub.com
SAGE

Abstract

An initial standardized approach for combining neuropsychological and neurophysiological measures in order to assess the neurocognitive effects of drugs in groups and individuals is introduced. Its application is illustrated with sedatives, antiepileptic drugs, psychostimulants, antihistamines, and intoxicants. Task performance, electroencephalography, and evoked potential measures during computerized attention and memory testing that are most sensitive to drug effects are identified in a sample population and then applied to individuals. In six example exploratory studies, drug effects were detected with an average area under curve (AUC) of 0.97 ($p < 0.0001$; 95% sensitivity, 96% specificity). In 10 example validation studies with other drugs and/or different subjects and populations, detection was strong in the eight studies with drugs and doses known to have significant neurocognitive effects (AUC 0.83, $p < 0.0001$; 82% sensitivity, 89% specificity), whereas no effect was detected in the two studies with drugs known to have faint neurocognitive effects (AUC 0.56, $p > 0.10$). Individual differences in response to different drugs with similar clinical uses, to varying doses of the same drug, and in pharmacodynamic response were then demonstrated. The significant ($p < 0.01$) increase in sensitivity and specificity of combined neuropsychological and neurophysiological measures compared with the former alone suggests that fewer subjects may be needed to assess the neurocognitive effects of drugs in future studies. The findings suggest that the concept of combining neuropsychological testing with simultaneous measures of neurophysiological function is worth further exploration.

Keywords

Brain function, cognition, drugs, individual differences

Introduction

Drug effects on the brain and cognitive processes can vary widely as a result of individual differences in biology and psychology (e.g. Cools et al., 2008; de Wit, 1998; Kabbaj, 2006; Park et al., 2007). An efficient and sensitive means to characterize such effects and individual differences therein would be helpful both during drug development and in clinical practice for optimizing individual patient treatment. Neuropsychological tests are usually employed to study drug-induced changes in cognitive function (Lezak et al., 2004; Spreen and Strauss, 1998). However, they lack direct measurements of brain activity. As a result such tests do not show the cause of poorer test performance (e.g. diminished alertness), nor do they detect compensatory efforts that can produce normal performance.

The electroencephalogram (EEG) is well known to be a sensitive indicator of drug action on the brain (Berger, 1931; Fink, 1984; Gloor, 1969; Hermann, 1982; Itil, 1974). Although not a three-dimensional imaging modality, unlike positron emission tomography (PET) or functional (f)MRI, EEGs can measure subsecond neurophysiological/neurocognitive processes, not visible to PET or fMRI, conveniently and comfortably during neuropsychological testing. In traditional 'pharmaco-EEG' studies of drug effects, subjects usually have been assessed resting with their eyes closed, obviating measurement of the interactions of neurological and cognitive processes. Over the past decade we have explored combining EEG measures with cognitive task performance measures, and found that the combined measures often were more sensitive to drug effects than either type of measure by itself (Gevins et al., 2002; Ilan et al., 2004; Meador et al., 2007; Smith et al., 2006). However, because these early studies employed a variety of paradigms, analytic methods, and EEG and task performance variables, and the results were not validated on new independent data, strong evidence for the general superiority of such a combined assessment was lacking. The earlier studies also lacked the means for determining how a drug affected

San Francisco Brain Research Institute and SAM Technology, San Francisco, CA, USA.

Corresponding author:
Alan Gevins, San Francisco Brain Research Institute and SAM Technology, 536 Waller Street, San Francisco, CA 94117, USA
Email: alan@sfbri.org

individual subjects within the populations tested. The current study aimed to extend past these limitations.

An initial approach to a standardized combined neuropsychological and neurophysiological assessment is introduced here. Examples of its applicability are shown in six exploratory studies of drugs with varying degrees of neurocognitive effects. Sedatives, antiepileptic drugs (AEDs), psychostimulants, an antihistamine, and intoxicants were used. Ten studies then illustrate assessment of construct validity by the degree to which the exploratory results generalize to new subjects and new drugs of the same types. Gauging retest reliability is demonstrated in a study that administered the same drug and dose to the same subjects on a different occasion. Examples of how to quantify individual differences are shown in steady-state responses to different drugs with the same indication for use, to varying doses of the same drug, and in pharmacodynamic response following a single dose. Finally, the sensitivity and specificity of detecting drug effects with and without EEG measures is compared. Based on these illustrative studies, we conclude that the concept of combined neuropsychological and neurophysiological testing is meritorious and worth further exploration.

Methods and materials

Test administration

Cognitive testing. Sustaining attention on an internal representation (Working Memory [WM]) is a fundamental cognitive process that is affected by many drugs and disorders (Chamberlain et al., 2006; Honey and Fletcher, 2006). A spatial WM task was therefore selected as the primary test for the initial battery. The well-known *n*-back WM task was used, in which subjects compared the location of dot stimuli on sequential trials (Gevins et al., 1990). Relatively more difficult (high load = comparing current stimulus location to that from two trials back) and relatively easier (low load = comparing current stimulus location to that from one trial back) versions of the task were administered by a computer program in blocks of 50 trials lasting approximately four minutes.

Of the many other cognitive functions that could be tested, verbal Episodic Memory (EM) was added to the initial battery because it is affected by a variety of drugs and diseases, such as Alzheimer's dementia (Belleville et al., 2008; Dickerson and Sperling, 2009; Fernandez-Serrano et al., 2010). A standard verbal EM task was used in which a list of words was first presented (encoding) by a computer program, followed by a distraction period, followed by testing for recognition of the previously presented words (Sanquist et al., 1980). During encoding, subjects indicated whether each sequentially presented word had one or two syllables by pressing the left or right mouse button. During recognition, they similarly indicated whether they had seen each word during the encoding phase (24 targets and 24 foils). In the seven studies that used the EM task (marked by a '+' in Table 1), the EM task was presented two times, with the same target words and different foils on each presentation. The WM task was administered between the word encoding and recognition phases, serving as the distracter (see Ilan et al., 2005; Meador et al., 2007 for details).

Electroencephalography. EEGs were recorded as subjects performed high and low load WM tasks (~8 min), during eyes closed (EC) and eyes open (EO) resting conditions to assess alertness (1.5 min each), and during the EM task when it was included in the battery (~6 min). A customized headset was used with head-mounted EEG amplifiers and disposable, solid-hydrogel electrodes placed over standard 10–20 System locations. These included bilateral and midline

Table 1. The combined assessment method was used for 11 drugs, in 6 exploratory studies (indicated by **bold** type; carbamazepine and topiramate were used together in one study), and 10 validation studies (regular type; for validation studies using the same drug as the exploratory study, the doses, number of subjects, age range, and peak interval for the validation studies are in reglular type and have different citation superscripts)[a]

Drug	Dose	No. of subjects	Age range (years)	Peak interval (s)
Lorazepam[10]+	**2 mg**	**10M/4F**	**20–41**	**90**
Alprazolam[2]	1 mg	4M/5F	21–35	90
Carbamazepine[6]	**mid-range therapeutic level**	**11M/17F**	**18–51**	**1 month steady state**
Topiramate[7]	**300 mg**	**12M/17F**	**22–58**	**1 month steady state**
Lamotrigine[7]	300 mg	12M/17F	22–58	1 month steady state
Levetiracetam[6]	2000 mg	11M/17F	18–51	1 month steady state
Methylphenidate[9]	**5, 10, 15, 20 mg**	**11M/2F**	**8–18**	**1 week steady state**
Caffeine[1,8]+	200mg[1], 250mg[8]	8M/8F[1], 9M/10F[8]	21–32[1], 21–55[8]	90[1], 30[8]
Diphenhydramine[1,8]+,[3]+	**50 mg[1,8,3]**	**8M/8F[1]**, 9M/10F[8], 6M/6F[3]	**21–32[1]**, 21–55[8], 62–75[3]	**150[1]**, (75 & 165)[8], 150[3]
Alcohol[1,8]+	**88 g/kg[1,8]**	**7M/8F[1]**, 8M/9F[8]	**21–32[1]**, 21–55[8]	**90[1]**, 75[8]
Marijuana[4]+,[5]+	**1.8 & 3.6% THC[4]**, 1.8 & 3.9% THC[5]	**11M/11F[4]**, 12M/10F[5]	**21–45[4]**, 22–37[5]	**20[4]**, 15[5]

[a]Studies with a '+' in the drug column employed a task battery that included a verbal episodic memory (EM) task in addition to the working memory (WM) task.
M = male, F = female. THC = tetrahydrocarbinol.
Peak intervals for acute studies are minutes after drug ingestion.
[1]Gevins et al. (2002); [2]McEvoy et al. (2001); [3]McEvoy et al. (2006); [4]Ilan et al. (2005); [5]Hart et al. 2010; [6]Meador et al. (2007); [7]Smith et al. (2006); [8]Gevins et al. in preparation-a; [9]Gevins et al. in preparation-b; [10]Meador et al. 2010.

dorsolateral prefrontal, midline sensorimotor, lateral superior parietal, and midline parieto-occipital cortical areas, referenced to digitally linked mastoids (Gevins et al., 2009). These locations were selected on the basis of cognitive EEG studies with 40 or 100 electrodes (Gevins et al., 1996, 1997, 2002). Vertical and horizontal eye movements were monitored with electrodes placed above and lateral to each eye. Signals were sampled at 128 Hz and band-pass filtered from 0.1 to 35 Hz. In three older studies (Table 1 – alcohol, caffeine, and diphenhydramine with young adults), EEGs were recorded from 40 electrodes, including the aforementioned sites, using a stretchable cap with gel-filled electrodes; signals were sampled at 256 Hz and band-pass filtered from 0.01 to 100 Hz. Data from these studies were digitally filtered and down-sampled to match the newer studies.

Data analysis

EEG and evoked potential analysis. Automatic detection and removal of artifacts due to eye movements and blinks, scalp muscle activity, head and body movements and bad electrode contacts (Du et al., 1994) was followed by visual inspection of all decontaminated and raw data. Power spectral estimates were computed on the decontaminated EEG data by averaging 2-second periodograms over the task or resting condition and combining them into several frequency band variables, as described below. Averaged evoked potentials (EPs) were computed across the interval from 0.5 second before to 1 second after the task stimulus. (The traditional neurological term 'evoked potential' is used herein, rather than the psychophysiological term 'event related potential', because the former is more generally known by non-specialists.)

A candidate set of six variables, or eight when the EM task was administered, was then computed from the task performance, EEG, and EP variables to characterize drug effects (Table 2). These variables were determined from prior research to be sensitive to drugs and confirmed by examination of the performance, EEG spectra, and EP averages. The two candidate task performance variables were divided for accuracy by reaction time measures from the WM and EM tasks, such that the variables were larger when responses were faster or more accurate (Gevins et al., 2002). The WM variables were averaged over the low load and high load versions of the task. The two candidate EP variables were mean slow wave amplitude (450–700 ms post-stimulus) during the WM and EM tasks, a component particularly sensitive to attention, memory load, and drugs (Garcia-Larrea and Cezanne-Bert, 1998; Ilan et al., 2004; Meador et al., 2007). The four candidate EEG power spectral variables included two measures especially sensitive to drowsiness: difference in alpha-band power between EC and EO resting conditions, and standard deviation of 2–6 Hz power in the EO resting condition (Hagiwara et al., 1997; Makeig and Jung, 1995; Oken and Salinsky, 1992), and two measures sensitive to both alertness and the systemic effects of drugs on mass neocortical synchronization – that is, the degree to which the drug is neuroactive: 2–6 Hz power and at 13–18 Hz power, averaged over all task and EO resting conditions (Clarke et al., 2008; Ilan and Gevins, 2001; Keane et al., 2007). These candidate variables were used in all six exploratory studies reported below. The variables were computed either after each drug had been taken repeatedly for a number of days to achieve a steady-state response or, in the case of single-dose studies, during the post-drug testing interval most closely matching the time of published peak level in plasma for each drug, and during a comparable placebo or non-drug interval.

Multivariate analysis of drug effects. The effects of a drug on a group of subjects were characterized first in exploratory analyses, as illustrated in Figure 1. Multivariate divergence analysis, a simple type of discriminant analysis, was used to find the subset of variables from the set of six or eight candidates that had the largest divergence between the drug and placebo/non-drug data sets (Gevins and Smith, 2006; Smith et al., 2001; Tou and Gonzalez, 1974). The divergence analysis considered all possible subsets of up to three variables and chose the subset of variables that, considered together as a group, best recognized a drug. It did not necessarily choose individual variables with statistically significant differences between drug and placebo/non-drug conditions, if such variables were highly correlated with each other.

Table 2. For each of the six exploratory studies, the same six candidate variables were submitted to the divergence analysis. An additional two candidate variables were considered when the episodic memory task was administered (studies indicated by '+' in Table 1). The analysis then selected the subset of up to three variables that best detected the effect of that drug. The descriptive terms in the 'Most sensitive to' column are simplifications for convenience (see the Discussion)

Candidate variable	Abbreviation	Most sensitive to:
Working Memory Accuracy and Reaction Time	WM performance	Performance
Episodic Memory Accuracy and Reaction Time	EM performance	Performance
WM Slow Wave Evoked Potential Amplitude	WM slow wave amplitude	Sustained Focused Attention
EM Slow Wave Evoked Potential Amplitude	EM slow wave amplitude	Episodic memory
EEG Alpha Reduction From Eyes Closed to Eyes Open	EC–EO alpha reduction	Alertness
Standard Deviation of 2-6 Hz EEG Power, Eyes Open	SD 2–6 Hz power	Alertness
2–6 Hz EEG Power, All Conditions Except Eyes Closed	2–6 Hz power	Neuroactivity
13–18 Hz EEG Power, All Conditions Except Eyes Closed	13–18 Hz power	Neuroactivity

Raw variable values were standardized within subjects, across conditions, so that all variables were on the same scale of standard deviation units.

For each new subject, the difference between the subject's value on each variable and the corresponding mean from the drug condition of the exploratory analysis group was computed. This difference for each variable was then weighted, based on the inverse of the variance of that variable in the exploratory data set and summed into a distance measure. Such measures were computed for each subject in all drug and dosage conditions. Values of the distance measures within each condition were then scaled and reported as a 'score' such that the mean scores in the exploratory analysis groups were −0.5 in the drug condition and +0.5 in the placebo/non-drug condition.

How strongly an individual subject's performance and brain function were affected by a particular drug and dose was quantified by the score for that subject. A relatively negative score in the drug condition indicated that the individual was affected by the drug in a similar manner to the exploratory analysis group. Whether this was considered deleterious or beneficial was based on how the drug affected the performance and EEG variables. A score near zero indicated that the individual evidenced little difference on the key variables between the drug and placebo/non-drug conditions, and a relatively positive score in the drug condition indicated that the individual had variable values that were more typical of the placebo/non-drug condition. Data from new subjects and/or other drugs or dosages collected in validation studies were likewise combined into scores for drug and

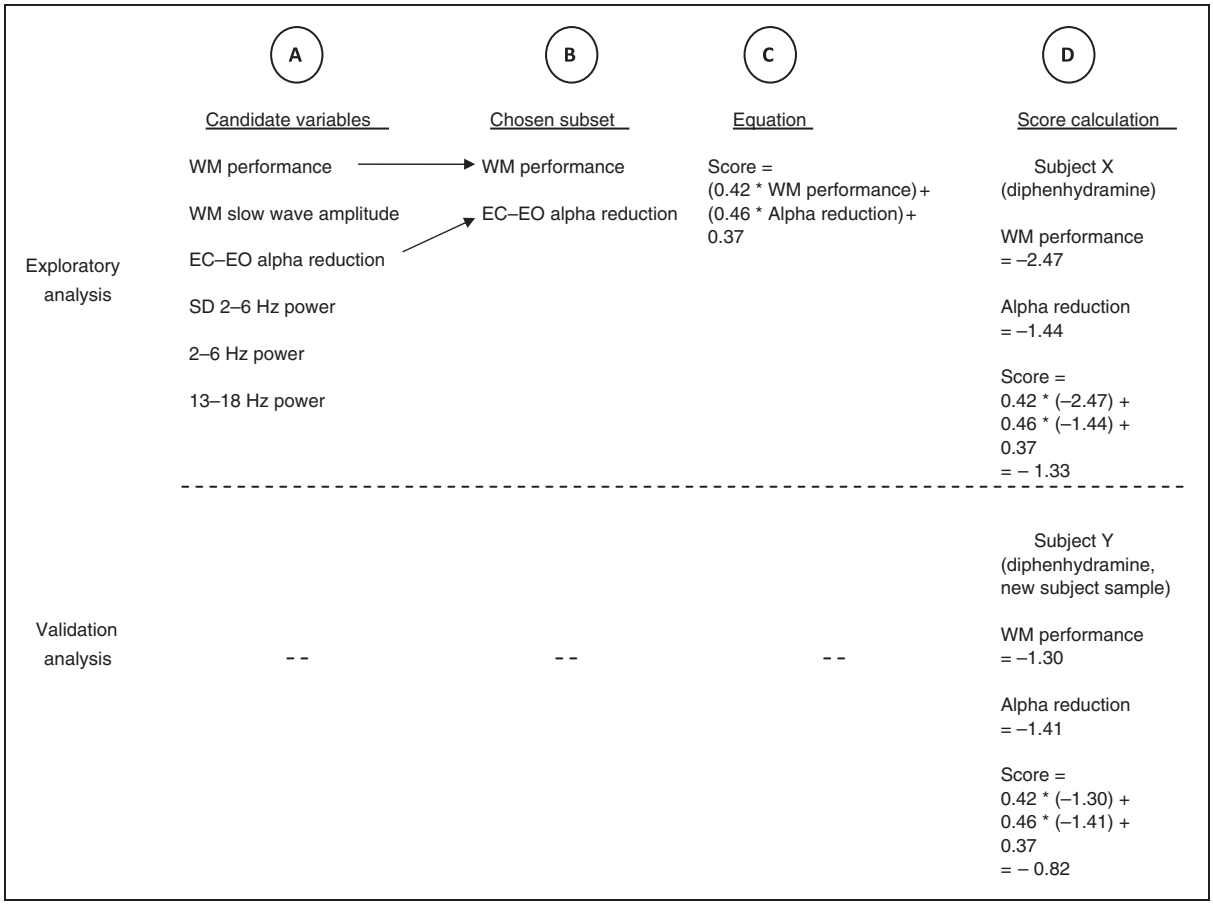

Figure 1. A simplified schematic illustration of how the divergence algorithm was used to create a single score characterizing a subject's response to a drug – in this example, diphenhydramine. *Top:* Exploratory Analysis. A. The divergence analysis considered all possible subsets of up to three variables from a set of six candidate variables. B. That subset was chosen that maximized the separation (divergence) between the diphenhydramine and placebo conditions in the exploratory data set. C. The variables in the chosen subset were weighted according to their contribution to the divergence, and a constant was added. The resulting scores were scaled so that the mean was −0.5 in the diphenhydramine condition and +0.5 in the placebo condition. D. Values of the chosen variables (in standard deviation units) from each subject and condition were entered into the equation to calculate scores characterizing subjects' neurocognitive responses to diphenhydramine or placebo. This approach allows easier interpretation of scores across applications. In the example, the subject's score of −1.33 means that this neurocognitive reaction to diphenhydramine was stronger (more negative) than that observed in the exploratory group as a whole (mean = −0.5). *Bottom*: Validation Analysis. D. In order to test the validity and generalizability of the equation for characterizing the neurocognitive effects of diphenhydramine, the equation generated in the exploratory analysis was applied to calculate scores for new subjects in drug and placebo conditions from other studies. In this example, a subject from a new sample of young adults had a score of −0.82 in the diphenhydramine condition.

placebo/non-drug conditions, using the same equation derived in the corresponding exploratory analysis. Validation results were compared with the exploratory results to determine whether the outcomes applied more generally. When applicable, pharmacodynamic changes over time were assessed by computing scores in the interval of the published peak level in plasma for each drug as well as in the prior and succeeding post-ingestion intervals.

Each array of subjects' scores from the drug and placebo/non-drug conditions of a study was used to generate a standard receiver–operator characteristic (ROC) curve (Hanley and McNeil, 1982). In order to compare the strength of the effects of different drugs and dosages, area under the curve (AUC), sensitivity (correctly identifying the drug condition), and specificity (correctly identifying the placebo/non-drug condition) were tabulated from the ROC curves. A drug with strong and consistent effects would tend to have separate clusters of scores for the drug and placebo/non-drug conditions, the easily discriminated scores between the two conditions leading to an AUC close to 1.0 with sensitivity and specificity both approaching 100%. A drug with weak and inconsistent effects would tend to yield highly overlapping scores in the drug and placebo/non-drug conditions, leading to an AUC close to 0.5 with sensitivity and specificity near 50%. AUC significance was calculated using the normal distribution from the z-ratio computed as the AUC minus 0.5 divided by the standard error for AUC = 0.5, the null hypothesis. Differences between two AUCs were assessed with a two-sample z-test. Differences between two arrays of AUCs were assessed with Wilcoxon signed-rank tests for smaller arrays and Student's t-tests for larger arrays.

Drugs and experimental designs

Six exploratory studies (Tables 1 and 3, bold type) and ten validation studies (Tables 1 and 3, regular type) illustrate how the combined assessment approach can be applied. All studies used randomized, double-blind, cross-over designs and were placebo-controlled except for the four antiepileptic drug (AED) studies, which used pre- and post-drug washout periods. Subjects received task training and practice during the first session of each study; these data were not analyzed. Exploratory analyses were performed for: (1) sedatives (lorazepam), (2) AEDs (carbamazepine and topiramate), (3) psychostimulants (methylphenidate), (4) antihistamines (diphenhydramine), (5) intoxicants (alcohol), and (6) intoxicants (marijuana). The exploratory results were then tested on other drugs (alprazolam for sedatives, lamotrigine and levetiracetam for AEDs, and caffeine for psychostimulants), or on different subjects taking the same drugs (diphenhydramine, alcohol, and marijuana). Retest reliability was illustrated by comparing results from the same subjects taking diphenhydramine on two occasions.

Acute drug effects were assessed by recording data prior to drug ingestion and again at the post-drug interval corresponding to the published t_{peak}, and comparing them to data from the corresponding placebo intervals. Steady-state effects of the four AEDs were assessed by testing subjects

Table 3. Area under the receiver–operator characteristic curve (AUC), sensitivity (percentage of subjects for whom the effect of the drug was correctly detected) and specificity (the percentage of subjects for whom the placebo or non-drug baseline was correctly identified) of the combined method for assessing effects of the drugs listed in Table 1[a]

Drug	AUC	Sensitivity	Specificity
Sedative (lorazepam)	**1.00*** **	**100%**	**100%**
Alprazolam, New Subjects	0.96**	100%	89%
Antiepileptic Drug - AED (Carbamazepine & Topiramate)	**0.98*** **	**93%**	**98%**
Lamotrigine	0.56	34%	100%
Levetiracetam	0.68†	75%	75%
Psychostimulant (Methylphenidate)	**0.96*** **	**100%**	**92%**
Caffeine, New Subjects	0.81*	88%	75%
Caffeine, Other New Subjects	0.84**	79%	89%
Antihistamine (Diphenhydramine)	**0.84***	**75%**	**94%**
Re-Test	0.89*	82%	100%
New Subjects	0.79*	74%	84%
Elderly Subjects	0.65	58%	92%
New Equation for Elderly	0.98***	100%	92%
Intoxicant (Alcohol)	**0.97*** **	**100%**	**93%**
New Subjects	0.88**	76%	100%
Intoxicant (Marijuana, Casual Users)	**1.00*** **	**100%**	**95%**
New Subjects, Frequent Users	0.94***	95%	82%

[a]The mean AUC for the six exploratory analyses (**bold** type) was 0.97 ($p < 0.0001$), with mean 95% sensitivity and 96% specificity. The successive indented rows (non-bold type) are the results of applying the corresponding exploratory analysis to new drugs of that type, to new subjects, or to a retest of the same drug on the same subjects. The mean AUC over the 10 validation studies was 0.78 ($p < 0.0001$, 76% sensitivity, 89% specificity). Strong effects were detected in each of the eight validation studies with drugs and doses known to have significant neurocognitive effects, whereas the detected effects were meager for the two studies with drugs known to have weak neurocognitive effects. (See the text for a discussion of the effect of diphenhydramine on elderly subjects.)

†$p < 0.05$, *$p < 0.01$, **$p < 0.001$, ***$p < 0.0001$.

after they had been taking a daily maintenance dose of the drug for 4 weeks, and comparing these tests to those administered before drug administration had begun as well as after a 4-week washout period. Effects of each of four doses of methylphenidate were assessed by testing subjects after the morning dose, after they had been taking a particular dose for a week, and comparing those tests to corresponding placebo data.

Subjects

All studies were reviewed and approved by an Institutional Review Board for human subjects research. In all but two studies, subjects were healthy adults between 18 and 55 years of age (Table 1). Subjects in the elderly diphenhydramine study were healthy and ranged from 62 to 75 years of age, and in the methylphenidate study subjects were patients from 8 to 18 years old who were being treated for attention-deficit/hyperactivity disorder (ADHD). Most studies had approximately equal numbers of males and females, the exceptions being the lorazepam and methylphenidate studies, which had a higher proportion of males. In the lorazepam, alprazolam, diphenhydramine, carbamazepine, levetiracetam, topiramate, and lamotrigine studies, subjects reported infrequent or no prior use of the study drug. Subjects in the alcohol and caffeine studies reported moderate usage of the study drugs (1–12 alcoholic drinks per week, 1–4 caffeinated drinks per day). In the marijuana (casual users) study, subjects reported at least 10 lifetime uses of marijuana, and in the marijuana (frequent users) study, subjects reported current daily use of marijuana.

Results

Exploratory analyses

The effects of the six types of drugs in the exploratory analyses were detected with a mean AUC of 0.97 ($p < 0.0001$) and mean 95% sensitivity and 96% specificity (Figure 2, black bars; Table 3, bold rows). The effect of the antihistamine diphenhydramine was notably weaker than that of the sedative, AED, psychostimulant, and intoxicant drug types. Task performance variables were used by the divergence analysis for five of the six drugs, the exception being alcohol (Table 4, bold entries). WM performance was impaired by the sedative, AEDs, and antihistamine and improved by the psychostimulant. EM performance was impaired by marijuana. EEG variables were used by the divergence analysis for all six drugs (Table 4, bold entries). Power in the 2–6 Hz band was reduced by marijuana, probably reflecting increased autonomic activation and alertness. Power in the 2–6 Hz band was increased by the older AEDs, an effect considered to be a marker of their neuroactivity (Mecarelli et al., 2004; Salinsky et al., 2004). Power in the 13–18 Hz band was increased by the sedative lorazepam, a characteristic marker of the neuroactivity of benzodiazepines (Knott, 2000). Power in the 13–18 Hz band decreased for methylphenidate, probably reflecting increased alertness, whereas the increased power in the 13–18 Hz band observed for alcohol was part of a characteristic neuroactive pattern of increasing EEG power that is seen across all frequency bands (Ehlers et al., 1989; Ilan and Gevins, 2001; Van Reen et al., 2006). A classic EEG marker of drowsiness (EC–EO alpha reduction) was chosen to detect

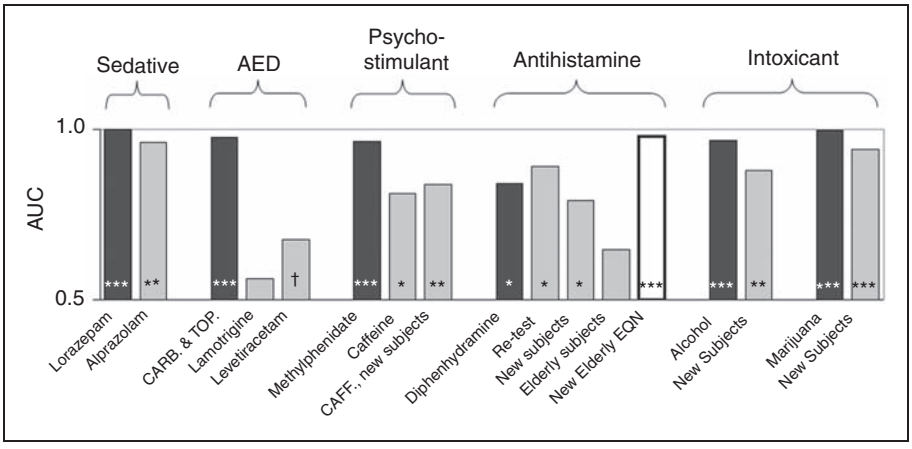

Figure 2. Area under the receiver–operator characteristic curve (AUC) for assessing cognitive and neurophysiological effects of the 11 drugs in the 16 studies. Black bars represent the exploratory studies and grey bars represent the validation results from applying the corresponding exploratory result to new drugs and/or new subjects, or a retest of the drug on the same subjects on a different occasion. Drugs known to have appreciable cognitive performance and neurophysiological effects were accurately detected by the exploratory equations, and the results generalized well in the validation studies to similar drugs and new subjects from the same population. The weaker effect for the two newer generation antiepileptic drugs (AEDs), lamotrigine and levetiracetam, is consistent with their well-documented milder neurocognitive side-effects relative to the older generation drugs, carbamazepine and topiramate, used to generate the exploratory equation. Diphenhydramine affected the elderly differently than the younger subjects whose tests were used to generate the exploratory equation, suggesting that they constituted a different population. Therefore a new exploratory equation was made for them (white bar). $^{\dagger}p < 0.05$, $*p < 0.01$, $**p < 0.001$, $***p < 0.0001$.

Table 4. Effects of the drugs on the candidate and final variables in each of the studies. Rows beginning with the drug type and names in **bold** type are the six exploratory studies; successive indented rows show the 10 validation results of applying the corresponding exploratory results to new drugs of that type, to new subjects, or to a retest of the same drug on the same subjects. An additional row shows the results from a new diphenhydramine equation for the elderly group[a]

Drug	2–6 Hz power	13–18 Hz power	EC–EO alpha reduction	SD 2-6 Hz Power	WM performance	WM slow wave amplitude	EM performance	EM slow wave amplitude
Sedative (lorazepam)	−0.01 ± 0.36	**1.13 ± 0.32****	−0.91 ± 0.34*	0.02 ± 0.37	**−1.67 ± 0.31*****	0.01 ± 0.38	−1.84 ± 0.26***	0.57 0.35
alprazolam	−1.79 ± 0.20***	**1.65 ± 0.13*****	−0.14 ± 0.34	−1.05 ± 0.39†	**−1.25 ± 0.32***	0.42 ± 0.47	N/A	N/A
Antiepileptic drug (AED) (carbamazepine & topiramate)	**1.17 ± 0.14*****	1.14 ± 0.13***	−0.78 ± 0.15***	0.87 ± 0.15***	**−1.09 ± 0.11*****	0.17 ± 0.15	N/A	N/A
Lamotrigine	**−0.16 ± 0.23**	−0.41 ± 0.23	−0.19 ± 0.26	0.14 ± 0.19	**−0.32 ± 0.18**	−0.37 ± 0.20	N/A	N/A
Levetiracetam	**−0.04 ± 0.23**	−0.54 ± 0.21†	−0.48 ± 0.17*	0.03 ± 0.20	**−0.44 ± 0.19**†	−0.09 ± 0.17	−0.28 ± 0.17	0.27 ± 0.21
Psychostimulant (methylphenidate)	−0.18 ± 0.30	**−0.46 ± 0.23**	−0.76 ± 0.43	−0.70 ± 0.32†	**0.62 ± 0.25**†	**0.70 ± 0.22***	N/A	N/A
Caffeine	−0.79 ± 0.34†	**−1.42 ± 0.37***	0.41 ± 0.37	−1.31 ± 0.25***	**0.31 ± 0.35**	**0.59 ± 0.41**	N/A	N/A
New caffeine subjects	−0.79 ± 0.33†	**−1.22 ± 0.28****	0.35 ± 0.38	−0.13 ± 0.37	**1.27 ± 0.41***	**0.47 ± 0.26**	0.44 ± 0.37	0.60 ± 0.33
Antihistamine (diphenhydramine)	0.51 ± 0.36	0.60 ± 0.35	**−1.18 ± 0.33***	0.87 ± 0.35†	**−1.12 ± 0.31***	0.43 ± 0.34	N/A	N/A
Re−test	0.81 ± 0.24†	0.99 ± 0.28†	**−1.18 ± 0.25***	0.96 ± 0.19*	**−0.79 ± 0.41**	1.67 ± 0.38	N/A	N/A
New subjects	0.38 ± 0.28	0.09 ± 0.24	**−0.72 ± 0.23***	0.54 ± 0.19†	**−0.37 ± 0.30**	0.51 ± 0.31	−0.57 ± 0.27†	0.24 ± 0.31
Elderly subjects	1.20 ± 0.35*	0.50 ± 0.40	**−0.12 ± 0.38**	1.76 ± 0.31**	**−0.40 ± 0.39**	0.14 ± 0.35	−0.32 ± 0.29	0.25 ± 0.40
New Equation for Elderly	1.20 ± 0.35*	0.50 ± 0.40	−0.12 ± 0.38	**1.76 ± 0.31****	−0.40 ± 0.39	**0.14 ± 0.35**	−0.32 ± 0.29	0.25 ± 0.40
Intoxicant (alcohol)	0.80 ± 0.39	**1.46 ± 0.14*****	−1.32 ± 0.34*	0.78 ± 0.31†	−0.30 ± 0.46	**0.35 ± 0.36**	N/A	N/A
New Subjects	−0.30 ± 0.40	**1.00 ± 0.21****	−0.07 ± 0.41	0.10 ± 0.37	−0.20 ± 0.41	**0.37 ± 0.38**	−0.64 ± 0.28†	−0.26 ± 0.35
Intoxicant (marijuana, casual users)	**−1.59 ± 0.16*****	−0.96 ± 0.27*	0.08 ± 0.26	−0.87 ± 0.30*	−0.65 ± 0.31	**−1.12 ± 0.25****	**−0.81 ± 0.25***	−1.00 ± 0.25**
New subjects, frequent users	**−1.46 ± 0.26*****	−0.96 ± 0.28*	0.03 ± 0.34	−0.77 ± 0.28†	−0.52 ± 0.24†	**−0.29 ± 0.22**	**−0.21 ± 0.22**	−0.56 ± 0.28

[a]Entries in each cell of the table are the standardized values (± standard error) showing a drug minus non-drug difference in the AED studies and a drug minus placebo difference in the other studies. **Bold** values are the final variables chosen by the divergence analysis to recognize each drug from its respective non-drug condition in the exploratory analysis that are then applied in the validation studies. A positive value indicates that the variable was larger in the drug than the respective non-drug condition.

WM = Working Memory task, EM = Episodic Memory task, SD = standard deviation, EC = Eyes-Closed, EO = Eyes-Open, N/A = Not applicable because the Episodic Memory task was not administered in the study.

†$p < 0.05$, *$p < 0.01$, **$p < 0.001$, ***$p < 0.0001$ refer to significance of difference between drug and non-drug condition for each variable as tested with a two-sided t-test, uncorrected for multiple comparisons to facilitate scanning the table across studies. (The Bonferroni-corrected significance cutoff for six comparisons is $p < 0.00833$ and for eight comparisons $p < 0.00625$).

the effects of diphenhydramine. WM slow wave amplitude, a marker of sustained focused attention, was chosen in three instances, being reduced by marijuana and increased by the psychostimulant and alcohol. EM slow wave amplitude was not chosen for any of the drugs.

Validation studies

Results of the six exploratory studies generalized well when applied to new data in the validation studies (Figure 2, gray bars; Table 3, regular rows), aside from elderly subjects taking diphenhydramine. The mean AUC over the 10 validation studies was 0.78 ($p < 0.0001$, 76% sensitivity, 89% specificity). Strong effects were detected in each of the eight validation studies with drugs and doses known to have significant neurocognitive effects (AUC 0.83, $p < 0.0001$; 82% sensitivity, 89% specificity), whereas no effects were detected for the two studies with drugs known to have weak neurocognitive effects (AUC 0.56, $p > 0.10$; 55% sensitivity, 88% specificity).

Sedatives. The exploratory analysis for the sedative lorazepam used impaired WM performance and increased 13–18 Hz power to recognize the effect of the drug with an AUC of 1.00 ($p < 0.0001$, 100% sensitivity, 100% specificity). When the exploratory result was applied to different subjects taking alprazolam, another sedating benzodiazepine with an intermediate course of action, the effects of alprazolam were recognized with an AUC of 0.96 ($p < 0.001$, 100% sensitivity, 89% specificity). This exploratory analysis generalized well to another drug in a different group of subjects because alprazolam had similar strong effects on cognitive performance and EEGs as did lorazepam (Table 4).

Antiepileptic drugs. When the result derived on the older AEDs carbamazepine and topiramate was applied to the same subjects taking the newer AEDs lamotrigine and levetiracetam, AUC decreased from 0.98 ($p < 0.0001$) to 0.56 ($p > 0.10$) and 0.68 ($p < 0.05$), respectively, reflecting the milder cognitive side-effects of these newer generation drugs. These studies are described in more detail in the section on individual differences.

Psychostimulants. The exploratory result for the psychostimulant methylphenidate generalized well to a demographically and clinically dissimilar population taking a different psychostimulant. In the exploratory study, 13 children and adolescents being treated for ADHD received daily doses of methylphenidate for a total of 5 weeks, according to a counterbalanced, crossover, double-blind design. For 1 week they each received placebo or 5, 10, 15, or 20 mg doses of methylphenidate. After being on a particular dose for a week, patients were tested 1–3 hours after ingesting the morning dose. The values of each of the six candidate EEG and performance variables were averaged over the four active doses and contrasted with placebo. Increased WM slow wave amplitude, improved WM performance, and decreased 13–18 Hz power were used to recognize the effects of methylphenidate (Table 4), with an AUC of 0.96 ($p < 0.0001$, 100% sensitivity, 92% specificity). Improved task performance and increased WM slow wave amplitude suggest improved attention (McEvoy et al., 1998), whereas decreased 13–18 Hz power reflects increased cortical activation and increased alertness (Barry et al., 2003; Gevins et al., 1979).

The methylphenidate exploratory result was then applied to healthy adults taking caffeine or placebo in separate sessions according to a randomized crossover design. In the study of Gevins et al. (2002), 16 subjects were given 200 mg caffeine. In a second study, 19 subjects took 250 mg caffeine. Data were analyzed near the time of peak plasma caffeine level (~1h post-drug ingestion; Liguori et al., 1997). The performance and EEG variables from the exploratory methylphenidate result were affected similarly by caffeine (Table 4), and the exploratory result thus generalized well, with AUCs of 0.81 ($p < 0.01$; 88% sensitivity, 75% specificity) and 0.84 ($p < 0.001$; 79% sensitivity, 89% specificity) for the studies with 200 mg and 250 mg of caffeine, respectively.

Antihistamine. When the same young adult subjects were re-tested with the sedating antihistamine diphenhydramine on a second occasion, the AUC was 0.89 ($p < 0.01$), compared with the 0.84 ($p < 0.01$) observed with the exploratory equation. The AUC was 0.79 ($p < 0.01$) for a new group of young adults who took diphenhydramine. Taken together, this might suggest that the exploratory results generalized well. However, the equation did not recognize the effects of diphenhydramine in a group of healthy elderly subjects (0.65 AUC, $p > 0.10$). Although EEG markers of decreased alertness were observed in both age groups after diphenhydramine (Table 4), they were manifest differently between the younger and older subjects. Younger subjects had less EC–EO alpha reduction after diphenhydramine, whereas the major change for elderly subjects was an increase in SD 2–6 Hz power, suggesting more varied alertness levels. Because the exploratory result on young adults was ineffective in characterizing the elderly adults' response to diphenhydramine, a new equation was derived on the elderly adult data. The new divergence analysis used increased SD 2–6 Hz power along with increased WM slow wave amplitude, and recognized the effects of diphenhydramine from placebo with an AUC of 0.98 ($p < 0.0001$, 100% sensitivity, 92% specificity; see 'New Equation for Elderly' in Tables 3 and 4, and 'New Elderly Eqn' in Figure 2).

Intoxicant, alcohol. When 16 subjects drank alcohol in quantities sufficient to raise blood/breath alcohol concentration to the legal driving limit (in California, 0.08g per 210L of breath), WM performance was not significantly affected. However, EEG power increased over a wide frequency range, consistent with previous findings (Table 4; Ehlers et al., 1989; Ilan and Gevins, 2001; Van Reen et al., 2006).

Two measures were used in the exploratory equation to recognize alcohol: increased 13–18 Hz power and increased WM slow wave amplitude. This resulted in an AUC of 0.97 ($p < 0.0001$, 100% sensitivity, 93% specificity).

The exploratory result was then applied to data from 19 new casual alcohol drinkers who were tested similarly in randomized, double-blind, crossover alcohol and placebo sessions after consuming a 500-mL drink containing either 0.88g/kg ethanol mixed in fruit juice, or a placebo drink containing 495 mL of fruit juice with 5 mL of alcohol floating on top to mimic the smell and taste of the active alcohol drink. The exploratory result generalized well to the new data resulting in recognition of alcohol's effect with an AUC of 0.88 ($p < 0.001$, 76% sensitivity, 100% specificity; see 'Alcohol, New Subjects' in Tables 3 and 4, and Figure 2).

Intoxicant, marijuana. The exploratory marijuana analysis used data from 22 infrequent users tested in a double-blind crossover design after smoking a cigarette with marijuana (1.8% or 3.6% Δ^9-tetrahydrocarbinol [THC]) or placebo (0% Δ^9-THC; Ilan et al., 2005). Twenty minutes after smoking, marijuana had numerous effects including decreased EM performance, EP amplitudes, and EEG power across a wide frequency range (Table 4), consistent with marijuana's initial stimulating effect and adverse effects on encoding and/or retrieval of verbal information (Ilan et al., 2004). Three measures (decreased 2–6 Hz power, decreased WM slow wave amplitude, and decreased EM performance) were selected by the divergence analysis with a resulting AUC of 1.00 ($p < 0.0001$, 100% sensitivity, 95% specificity). The validation study employed a similar design but assessed the effects of marijuana in 22 daily users (Hart et al., 2010). Marijuana had similar but generally weaker effects on the frequent users (Table 4). The exploratory result generalized well with an AUC of 0.94 ($p < 0.0001$, 95% sensitivity, 82% specificity; see 'Marijuana, New Subjects, Frequent Users' in Tables 3 and 4, and 'Marijuana, New Subjects' in Figure 2).

Assessing individual differences in response to drugs

Three studies illustrated the use of combining neurophysiological and task performance measures to examine the effect of a drug on each individual relative to a group of subjects. The first examined individual steady-state responses to traditional and newer AEDs, the second examined individual differences in neurocognitive pharmacodynamics of diphenhydramine, and the third examined individual dose–response variations to methylphenidate. In all cases, a score for each subject was computed by reference to the exploratory result as described above.

Individual differences in response to antiepileptic drugs. AEDs tend to decrease the incidence of propagation of abnormal epileptogenic discharges in the EEG, but their effects on excitatory and inhibitory neurotransmission can result in cognitive impairment. Two AEDs (topiramate and carbamazepine) that are known to have generally adverse neurocognitive effects (Mecarelli et al., 2001; Placidi et al., 2004; Salinsky et al., 2002) were used for the exploratory study, and two AEDs (lamotrigine and levetiracetam) with milder cognitive side-effect profiles (Marciani et al., 1999; Mecarelli et al., 2004) were used for the validation study. Twenty-nine healthy subjects were tested after reaching steady states on topiramate and lamotrigine (Smith et al., 2006), and 28 healthy subjects were tested after reaching steady states on carbamazepine and levetiracetam (Meador et al., 2007). In both studies, the drug conditions were compared to non-drug baselines taken before drug administration and after a 1-month post-drug washout period. The exploratory analysis distinguished carbamazepine and topiramate from the non-drug conditions with an AUC of 0.98 ($p < 0.0001$, 93% sensitivity, 98% specificity) using decreased WM performance and increased 2–6 Hz power (Table 4). The exploratory result did not generalize to lamotrigine (AUC 0.56, $p > 0.10$) and only weakly generalized to levetiracetam (AUC 0.68, $p < 0.05$; 75% sensitivity, 75% specificity). Consistent with clinical observations, these results suggest that lamotrigine and levetiracetam had milder cognitive and EEG effects than carbamazepine and topiramate.

Scores were generated for each of the 57 subjects after taking the study drugs for 1 month (Figure 3). Most individuals followed the group pattern and showed smaller neurocognitive effects in response to lamotrigine or levetiracetam than to topiramate or carbamazepine. However, in a number of instances subjects had a strong reaction to both drugs or even showed a stronger response to lamotrigine or levetiracetam. For instance, subject mlt8823's strong negative reaction to lamotrigine was comparable to his response to topiramate (Figure 3(A)). His WM performance was adversely affected and his 2–6 Hz power was increased by both drugs. Similarly, subject aed03 had comparable negative reactions to levetiracetam and carbamazepine with poorer WM performance and increased 2–6 Hz power relative to his non-drug baseline (Figure 3(B)). By contrast, several subjects (e.g. mlt8819, aed23, and aed01) were relatively unaffected by any of the AEDs, with no observed adverse effects on WM performance or 2–6 Hz power.

Individual differences in neurocognitive pharmacodynamics. The exploratory analysis was performed on 12 elderly subjects in the interval 150 min after they ingested 50 mg of diphenhydramine (corresponding with the published t_{peak}). Scores were computed for each individual at the peak interval and at intervals one hour before and one hour after (Figure 4). The group means (leftmost set of bars) suggest a classic U-shaped response curve with large negative cognitive and neurophysiological responses to diphenhydramine 150 min post-drug and responses at 90 and 210 min post-drug that were not significantly different from the post-placebo response. However, there were notable individual differences. Whereas some subjects, such as a01, a02, and a07, had the same pharmacodynamic pattern as the group, others differed. For instance, the peak response for subject a11 was largest during the 210-min interval, while the response of subject a12 occurred earlier during the 90-min interval. Perhaps of most clinical interest were the large pharmacodynamic variations after an individual's peak effect, with 4 of the 12 subjects having a persisting response at the 210-min interval.

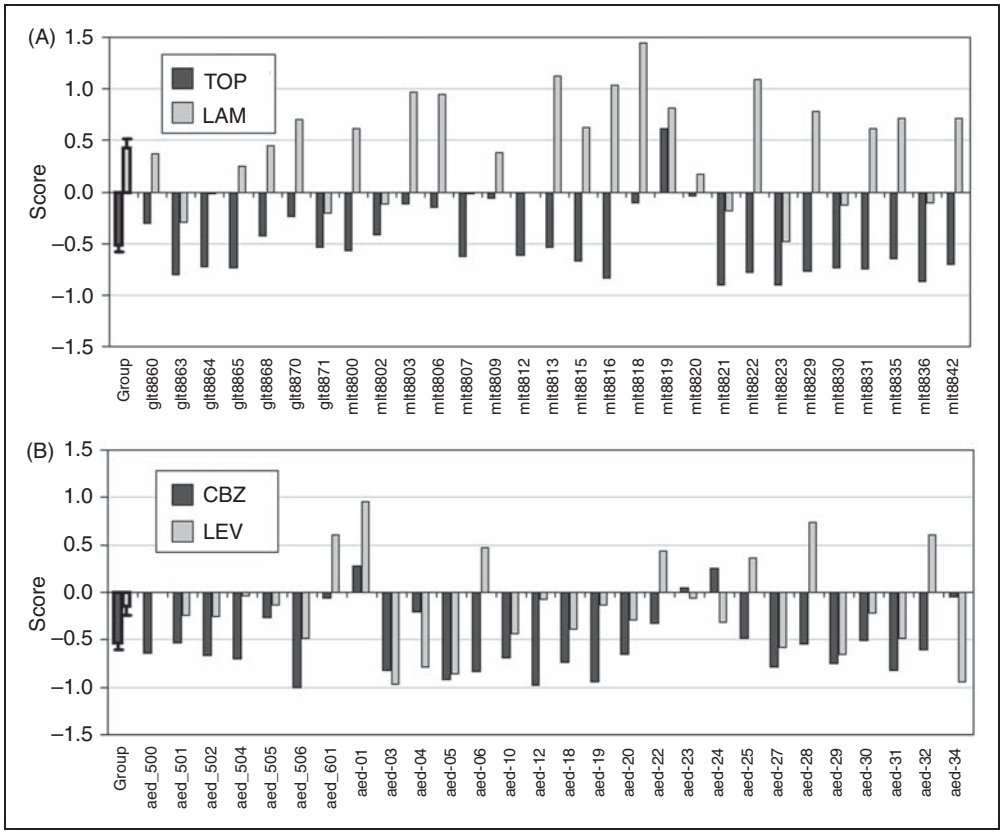

Figure 3. (A) Individual differences in response to steady-state levels of antiepileptic drugs (AEDs) are shown in the scores of 29 subjects who had been taking topiramate (black bars) or lamotrigine (grey bars) for 28 days. The scores were produced for each subject on each drug by comparing the values of the variables for that drug to the exploratory result that discriminated the two older generation AEDs, topiramate and carbamazepine, from non-drug conditions (Tables 3 and 4 and Figure 2). The means (and standard errors) for the groups of subjects are shown in the leftmost bars. A more negative score indicates a stronger cognitive and neurophysiological effect of the AED. Most individuals followed the group pattern and had milder cognitive performance and neurophysiological responses to the two newer generation AEDs, lamotrigine and levetiracetam. However, there are a number of exceptions in which subjects had a strong reaction to both drugs or even showed a stronger response to the newer generation drug. (B) Individual differences in response to steady-state levels of antiepileptic drugs (AEDs) are shown in the scores of 28 subjects who had been taking carbamazepine (black bars) or levetiracetam (grey bars) for 28 days. The means (and standard errors) for the groups of subjects are shown in the leftmost bars. The scores were produced in the same manner as is described in Figure 3(A).

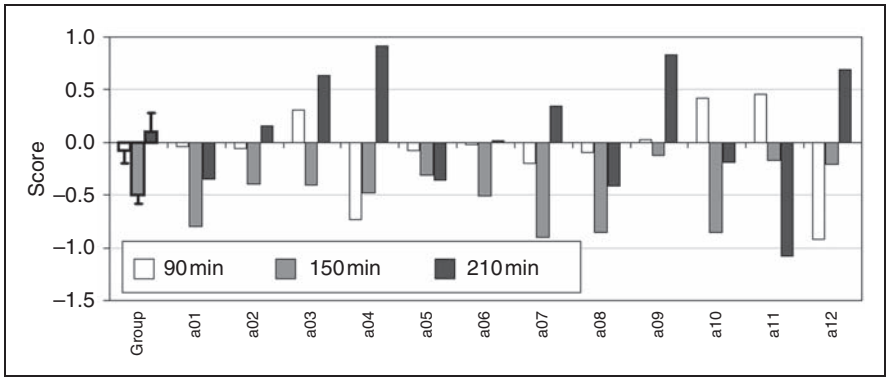

Figure 4. Individual differences in neurocognitive pharmacodynamics are shown by variations in scores of 12 elderly subjects 90 min (white bars), 150 min (grey bars), and 210 min (black bars) after taking 50 mg of diphenhydramine. A more negative score indicates a stronger cognitive and neurophysiological effect of diphenhydramine. The scores were produced by comparing each subject's variable values at each of the three measurement time points to the exploratory result that discriminated diphenhydramine at the time of peak effect (150 min post-drug) from placebo (Tables 3 and 4 and Figure 2). The group mean and standard error for each time point appear at left. Several subjects had their strongest response 60 min before or after the group peak response, and there were large pharmacodynamic variations after individuals' peak effects.

Individual differences in dose–response function. The exploratory analysis was derived on the average of all methylphenidate doses in 13 child and adolescent patients being treated for ADHD. Drug response scores were computed for each patient at each of four doses (5, 10, 15 and 20 mg; Figure 5). Out of the 13 patients, six had a peak response at 5 mg, two at 10 mg, two at 15 mg, and three at 20 mg. The lack of a linear relation between the scores and the dose is consistent with the idiosyncratic therapeutic dose–response relationship for treatment of ADHD with methylphenidate (Swanson and Volkow, 2002). For instance, patient mph006 had small responses at 5, 10, and 15 mg but a high score at 20 mg resulting from the combination of increased WM slow wave EP amplitude (6 μV larger than placebo) and better WM performance (reaction time [RT] 108 ms shorter). The largest score for patient mph005 was at 15 mg with a smaller response at 20 mg, primarily because WM slow wave amplitude increased more, relative to placebo, in the 15 mg condition than in the 20 mg condition. By contrast, patient mph009 had the largest score at 5 mg, a smaller score at 10 mg and a response similar to that in the placebo condition at 15 mg. At 5 mg WM performance was over 300 ms shorter and 10% more accurate and slow wave amplitude was 4 μV larger than in the placebo condition.

Value-added of EEG measures

In all the above studies, EEG measures had an important role in the recognition of drug effects. Across the six exploratory studies, AUC was 0.97 (95% sensitivity, 96% specificity) in detecting drug effects using a combination of EEG and task performance measures. This is significantly higher than the AUC of 0.87 (76% sensitivity, 87% specificity) achieved when the divergence analysis was run with an equivalent number (six or eight) of only task performance measures ($z = 2.1$, $p < 0.05$). Similarly, across all 16 studies listed in Table 1, a significant improvement in drug effect detection was obtained using a combination of EEG and task performance variables (0.86 AUC, 86% sensitivity, 91% specificity) compared with using an equivalent number of task performance measures by themselves (0.76 AUC, sensitivity 73%, specificity 77%), $t(15) = 3.5$, $p < 0.01$.

The marijuana analysis illustrates this effect. When the eight candidate EEG and performance variables were considered, the divergence analysis selected three final variables (2–6 Hz power, EM performance, and WM slow wave amplitude; Table 4) to distinguish marijuana from placebo data, resulting in an AUC of 1.00 ($p < 0.0001$, 100% sensitivity, 95% specificity). When eight task performance variables by themselves were considered (accuracy, mean RT, and standard deviation of RT in the low and high load versions of the WM task, and accuracy and a composite RT measure from the EM task), three variables were chosen (RT in the high load WM task, accuracy in the EM task, and standard deviation of RT in the low load WM task), resulting in an AUC of 0.78 ($p < 0.01$, 68% sensitivity, 77% specificity), significantly lower than the AUC from the combined analysis ($z = 3.13$, $p < 0.01$). In addition, the ability of the exploratory analysis to generalize to new data decreased when EEG variables were not used. When the equation with both types of measures was applied to new subjects, marijuana was detected with an AUC of 0.94 ($p < 0.0001$, 95% sensitivity, 82% specificity; Figure 2 and Table 3). The generalization was significantly lower ($z = 2.77$, $p < 0.01$) when the equation using only task performance measures was applied to the new subjects, as AUC was 0.70 ($p < 0.05$, 95% sensitivity, 45% specificity).

Discussion

An initial version of a standardized combined assessment that efficiently measures the neurocognitive effects of drugs on

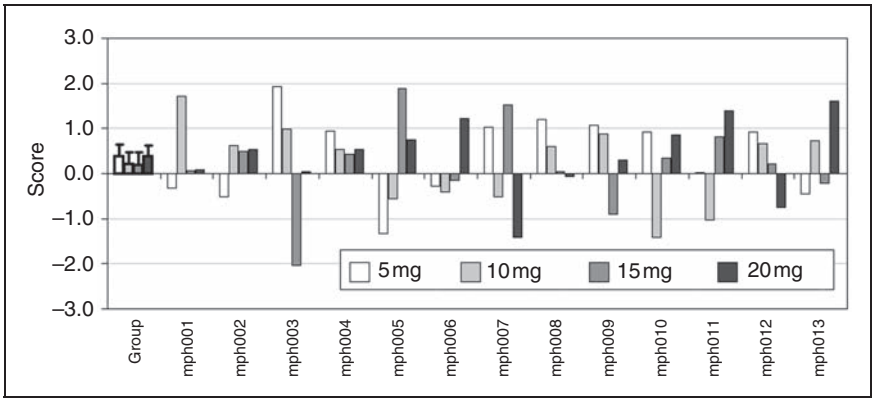

Figure 5. Individual differences in response to varying therapeutic doses of a psychostimulant are shown in the scores of 13 ADHD patients (ages 8–18years) after they had taken each of four different doses of methylphenidate for 1week: 5 mg (white bars), 10 mg (light grey bars), 15 mg (dark grey bars), and 20 mg (black bars). The scores were produced for each subject at each dose by comparing their variables to the exploratory result derived on the group to discriminate all doses of methylphenidate from placebo (Tables 3 and 4 and Figure 2). The signs of the scores in the figure have been arranged such that more positive scores reflect a stronger beneficial response to methylphenidate relative to placebo, including better working memory task performance and increased evoked potential and electroencephalogram markers of attention and alertness. Negative scores indicate responses similar to those in the placebo condition. The group means and standard errors for each dose appear at the left. The lack of a linear relation between the scores and the dose is consistent with the well-known idiosyncratic therapeutic dose–response relationship for methylphenidate.

groups and individuals was described. The method combines neuropsychological testing with simultaneous EEG recording. In exploratory studies, a multivariate divergence analysis selected a combination of up to three task performance and EEG/evoked potential variables that best recognized the effect of a drug in a group of subjects. The exploratory results were then tested in validation studies with new data from the same or new subjects with the same or another drug or dose. Individual differences in response to a drug were quantified by computing a score for each subject by reference to the exploratory group result.

Exploratory studies were illustrated with sedatives, AEDs, psychostimulants, antihistamines, and intoxicants. The effects of these drugs were recognized with an AUC of 0.97 ($p < 0.0001$, 95% sensitivity, 96% specificity). The 10 validation studies had an AUC of 0.78 ($p < 0.0001$; 76% sensitivity, 89% specificity). Strong effects were detected in each of the eight validation studies with drugs and doses known to have significant neurocognitive effects (AUC 0.83, $p < 0.0001$; 82% sensitivity, 89% specificity), such as alprazolam, whereas no effects were detected for the two studies with drugs known to have slight neurocognitive effects (AUC 0.56, $p > 0.10$; 55% sensitivity, 88% specificity), such as lamotrigine. The combined assessment demonstrated good discriminative validity and good face validity in these illustrative examples. Good retest reliability was found when the same subjects took the same drug on another occasion, for example, retest sensitivity and specificity for diphenhydramine were 82% and 100%, respectively (AUC 0.89, $p < 0.01$). A wide variation in individual subjects' responses to drugs was observed, including in drugs with the same indication for use (e.g. the AEDs), in different doses of the same drug (e.g. methylphenidate), and in pharmacodynamic response (e.g. diphenhydramine effects in the elderly).

Use of the combined assessment to determine whether subjects from different populations had the same response to a drug or drug class as the exploratory analysis population was also illustrated. For instance, two groups of healthy adults had similar responses to caffeine as a group of children and adolescents with ADHD did to the psychostimulant methylphenidate. In contrast, the response of healthy elderly adults to the sedating antihistamine diphenhydramine differed from that of the healthy young adults, as the diminished alertness of the elderly subjects was more variable.

Results of the AED studies illustrated how the combined assessment might be used to compare the cognitive performance and neurophysiological effects of new drugs with the same indication of use as established drugs, and to quantify individual variations in drug response. Further, results from individual dose responses to methylphenidate pointed to how such an assessment might someday be used to provide objective evidence about which dose of a given drug is most suitable for an individual patient. Because the final variables in the exploratory psychostimulant analysis reflected generally beneficial effects of methylphenidate, the dose resulting in the largest score for a particular patient might be considered an 'optimal neurocognitive dose'. If used in conjunction with other information, such as parent and teacher ratings, such scores might help a physician more quickly determine which dose would provide the largest benefit without overmedicating the patient.

The results also demonstrated the benefit of adding neurophysiological measures to cognitive task performance measures for characterizing drug effects. Because sensitivity and specificity increased significantly when both performance and neurophysiological measures were used relative to using only performance measures, fewer subjects may be required to assess drug effects when both types of measures are used. For instance, if a hypothetical drug had an effect as strong as the average observed over the 16 studies reported here, only 20 subjects would be needed to detect its effect (with 90% power and a 1% typeI error rate) if both EEG and task performance measures were used, as compared with 43 subjects using only the task performance measures. Whether or not achieving an improvement in sensitivity and reduction in the number of subjects required would be worth the extra effort and expense of obtaining and analyzing EEGs would depend on the particulars of the drug and population being studied.

The combined assessment has shown promising results with prescriptions, over-the-counter drugs, and recreational drugs in measuring chronic steady-state and single-dose acute drug effects and in tracking pharmacodynamic responses over time. However, the particular combined assessment and the illustrative studies presented here merely serve to demonstrate some of the possible applications, and the added value of combining neuropsychological task performance and simultaneous neurophysiological measures. The current results have many limitations to be addressed in future work. For instance, the neuropsychological battery only tested spatial working memory and verbal episodic recognition memory, and did not span a broader range of cognitive functions. Likewise, the candidate set of cognitive performance and EEG variables, although effective for the drugs used in the exploratory and validation studies, may well not be the best ones for other types of drugs or populations.

Because the illustrative studies had relatively small numbers of subjects, the number of candidate variables was limited to six or eight and the number of final variables was limited to a maximum of three to avoid over-fitting the data (Gevins, 1980). Future research or clinical applications would benefit from testing a larger number of subjects, including patients. As a convenience, the candidate EEG variables were labeled as being mostly sensitive to alertness, neuroactivity, sustained attention, and EM but this is an oversimplification. In fact, the relations are more complex, as alertness influences attention and the neuroactive effect of a drug may be associated with changes in alertness. Such subtleties of interpretation are a consideration not only for EEG measures but for all measures of functional activity in the brain including PET and fMRI. Simple and well-established EEG power spectral and EP component amplitude measures were used, but more complex ones (e.g. wavelet coefficients, independent components analysis, time-varying spectra) might prove more sensitive. Multivariate divergence analysis was used because of its simplicity and ease of interpretation compared with other more complex multidimensional variable subset selection and pattern classification algorithms

(e.g. genetic algorithms, neural networks, support vector classifiers), but such techniques may be more effective. Finally, other neuroimaging modalities such as PET or fMRI could be used instead of EEG.

With further research followed by prospective clinical trials, a combined assessment might someday prove to be a valuable tool in improving individualized medicine by helping physicians match an individual patient with a particular drug and dose that would maximize the neurocognitive therapeutic benefit while minimizing negative side-effects. In the meantime, such methods may prove useful with compounds under development by providing a more sensitive measure of the neurocognitive effects and side-effects of drugs in groups and in individual subjects and patients.

Acknowledgements

We gratefully acknowledge and appreciate the efforts of our scientist collaborators and their colleagues in whose laboratories much of the data were collected, including: Dr Kimford Meador and Dr David Loring (AEDs) Department of Neurology, Emory University, Atlanta, GA; Dr Harriet de Wit, University of Chicago (marijuana); Dr Carl Hart, Columbia University (marijuana); Dr Sandra Loo and Dr James McGough, Department of Psychiatry, UCLA (methylphenidate); and Dr Philip Leese, Quintiles and Dr Christian Otoul, UCB Pharma (lorazepam). We also acknowledge and appreciate the efforts of our colleagues at the San Francisco Brain Research Institute and SAM Technology for their contributions to the studies using diphenhydramine, caffeine, alcohol, marijuana and alprazolam, and their other vital scientific and engineering contributions to development of the method, including especially Dr Michael Smith, Dr Linda McEvoy, Dr Harrison Leong, Dr Robert Du, Mr Patrick Sullivan, Mr Guogang Hu, Mr Mathieu Herbette, Mr Gordon Abraham, Mr Robert Howard, Mr Art Sandoval, Dr Michele Lazzarra, Dr Mayra Padilla, Ms Brynn Evans, Ms Kemi Role, Ms Emilie Schwager, Ms Jamie Elmasu, and Mr Zachary Davis. We thank Dr Max Fink, Dr Peter Irwin, Dr Enoch Callaway, and Dr Kimford Meador for critical reviews of drafts of this paper.

Funding

This research was supported by grants from The National Institute of Neurological Diseases and Strokes, The National Institute of Mental Health, The National Institute of Alcoholism and Alcohol Abuse, The National Institute of Drug Abuse, The National Heart Lung and Blood Institute, The National Institute of Aging, and The National Institute of Child Health and Human Development (Alan Gevins, Principal Investigator).

Conflict of interest statement

No industry support was received to perform the work described herein. Underlying studies of levetiracetam and carbamazepine, and lamotrigine and topiramate were performed by Dr Kimford Meador and were supported by UCB Pharma and Glaxo-Smith-Kline, respectively. The EEG and associated cognitive testing of subjects in those studies, and the analyses of the data, were supported by Dr Gevins' NIH grants. Lorazepam data came from a study of two AEDs sponsored by UCB Pharma and performed by Quintiles Inc. using testing procedures like that reported here; analyses of the lorazepam data reported herein were supported by Dr Gevins' NIH grants.

References

Barry RJ, Johnstone SJ and Clarke AR (2003) A review of electrophysiology in attention-deficit/hyperactivity disorder. II. Event-related potentials. *Clin Neurophysiol* 114: 184–198.

Belleville S, Sylvain-Roy S, de Boysson C and Menard MC (2008) Characterizing the memory changes in persons with mild cognitive impairment. *Prog Brain Res* 169: 365–375.

Berger H (1931) Uber das Elektroenkephalogramm des Menschen. Drittem Mitteilung. *Archiv Psychiatrie Nervenkrankheiten* 94: 16–60.

Chamberlain SR, Muller U, Robbins TW and Sahakian BJ (2006) Neuropharmacological modulation of cognition. *Curr Opin Neurol* 19: 607–612.

Clarke AR, Barry RJ, McCarthy R, Selikowitz M and Johnstone SJ (2008) Effects of imipramine hydrochloride on the EEG of children with Attention-Deficit/Hyperactivity Disorder who are non-responsive to stimulants. *Int J Psychophysiol* 68: 186–192.

Cools R, Gibbs SE, Miyakawa A, Jagust W and D'Esposito M (2008) Working memory capacity predicts dopamine synthesis capacity in the human striatum. *J Neurosci* 28: 1208–1212.

de Wit H (1998) Individual differences in acute effects of drugs in humans: their relevance to risk for abuse. *NIDA Res Monogr* 169: 176–187.

Dickerson BC and Sperling RA (2009) Large-scale functional brain network abnormalities in Alzheimer's disease: insights from functional neuroimaging. *Behav Neurol* 21: 63–75.

Du W, Leong HM and Gevins AS (1994) Ocular artifact minimization by adaptive filtering. In: *Proceedings of the Seventh IEEE SP Workshop on Statistical Signal and Array Processing*, Quebec City, Canada, pp. 433–436.

Ehlers CL, Wall TL and Schuckit MA (1989) EEG spectral characteristics following ethanol administration in young men. *Electroencephalogr Clin Neurophysiol* 73: 179–187.

Fernandez-Serrano MJ, Perez-Garcia M and Verdejo-Garcia A (2010) What are the specific vs generalized effects of drugs of abuse on neuropsychological performance? *Neurosci Biobehav Rev* 00: 0–0 (accessed 1 November 2010).

Fink M (1984) Pharmacoelectroencephalography: a note on its history. *Neuropsychobiology* 12: 173–178.

Garcia-Larrea L and Cezanne-Bert G (1998) P3, positive slow wave and working memory load: a study on the functional correlates of slow wave activity. *Electroencephalogr Clin Neurophysiol* 108: 260–273.

Gevins A, Chan C and Ilan A (in preparation-b) Cognitive and neurophysiological effects of methylphenidate in children with ADHD.

Gevins A, Howard R and Sandoval A (2009) EEG Electrode Headset. U.S. Patent No. 7551952.

Gevins A, Ilan A, Chan C and Baum C (in preparation-a) Combined cognitive and neurophysiological effects of caffeine, diphenhydramine, and alcohol.

Gevins A, Smith ME, McEvoy L and Yu D (1997) High-resolution EEG mapping of cortical activation related to working memory: effects of task difficulty, type of processing, and practice. *Cerebral Cortex* 7: 374–385.

Gevins A, Smith ME and McEvoy LK (2002) Tracking the cognitive pharmacodynamics of psychoactive substances with combinations of behavioral and neurophysiological measures. *Neuropsychopharmacology* 26: 27–39.

Gevins AS, Bressler SL, Cutillo BA, et al. (1990) Effects of prolonged mental work on functional brain topography. *Electroencephalogr Clin Neurophysiol* 76: 339–350.

Gevins AS and Smith ME (2006) EEG in neuroergonomics. In: Parasuraman R and Rizzo M (eds) *Neuroergonomics: The*

Brain at Work. Oxford and New York: Oxford University Press, pp. 15–31.

Gevins AS, Smith ME, Le J, et al. (1996) High resolution evoked potential imaging of the cortical dynamics of human working memory. *Electroencephalogr Clin Neurophysiol* 98: 327–348.

Gevins AS (1980) Pattern recognition of brain electrical potentials. *IEEE Trans Patt Anal and Mach Intell PAMI-2* 5: 383–404.

Gevins AS, Zeitlin GM, Doyle JC, et al. (1979) Electroencephalogram correlates of higher cortical functions. *Science* 203: 665–668.

Gloor P (1969) Hans Berger on the electroencephalogram of man. *Electroencephalogr Clin Neurophysiol* Suppl. 28 (translated edition): 1–36.

Hagiwara H, Araki K, Michimori A and Saito M (1997) [A study on the quantitative evaluation method of human alertness and its application]. *Seishin Shinkeigaku Zasshi* 99: 23–34.

Hanley JA and McNeil BJ (1982) The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology* 143: 29–36.

Hart CL, Ilan AB, Gevins A, Gunderson EW, Role K, Colley J and Foltin RW (2010) Neurophysiological and cognitive effects of smoked marijuana in frequent users. *Pharmacol Biochem Beh* 96: 333–341.

Hermann WM (1982) *Electroencephalography in Drug Research*. Stuttgart: Gustav Fischer.

Honey GD and Fletcher PC (2006) Investigating principles of human brain function underlying working memory: what insights from schizophrenia? *Neuroscience* 139: 59–71.

Ilan AB and Gevins A (2001) Prolonged neurophysiological effects of cumulative wine drinking. *Alcohol* 25(3): 137–152.

Ilan AB, Gevins A, Coleman M, ElSohly MA and de Wit H (2005) Neurophysiological and subjective profile of marijuana with varying concentrations of cannabinoids. *Behav Pharmacol* 16: 487–496.

Ilan AB, Smith ME and Gevins A (2004) Effects of marijuana on neurophysiological signals of working and episodic memory. *Psychopharmacology* 176: 214–222.

Itil TM (1974) *Modern problems of pharmacopsychiatry: Psychotropic Drugs and the Human EEG*. Vol 8, New York: Karger.

Kabbaj M (2006) Individual differences in vulnerability to drug abuse: the high responders/low responders model. *CNS Neurol Disord Drug Targets* 5: 513–520.

Keane MA, James JE and Hogan MJ (2007) Effects of dietary caffeine on topographic EEG after controlling for withdrawal and withdrawal reversal. *Neuropsychobiology* 56(4): 197–207.

Knott VJ (2000) Quantitative EEG methods and measures in human psychopharmacological research. *Hum Psychopharmacol* 15: 479–498.

Lezak MD, Howieson DB and Loring DW (2004) *Neuropsychological Assessment*, 4th ed. New York: Oxford University Press.

Liguori A, Hughes JR and Grass JA (1997) Absorption and subjective effects of caffeine from coffee, cola and capsules. *Pharmacol Biochem Behav* 58: 721–726.

Makeig S and Jung TP (1995) Changes in alertness are a principal component of variance in the EEG spectrum. *NeuroReport* 7: 213–216.

Marciani MG, Spanedda F and Mattia D (1999) Neurophysiologic and neuropsychologic profiles of lamotrigine in epilepsy. *Clin Neuropharmacol* 22: 159–163.

McEvoy LK, Pellouchoud E, Smith ME, and Gevins A (2001). Neurophysiological signals of working memory in normal aging. *Cognitive Brain Research* 11: 363–376.

McEvoy LK, Smith ME, Fordyce M, Gevins A (2006) Characterizing impaired functional alertness from diphenhydramine in the elderly with performance and neurophysiologic measures. *Sleep* 29: 959–966.

McEvoy LK, Smith ME and Gevins A (1998) Dynamic cortical networks of verbal and spatial working memory: Effects of memory load and task practice. *Cerebral Cortex* 8: 563–574.

Meador KJ, Gevins A, Loring DW, et al. (2007) Neuropsychological and neurophysiologic effects of carbamazepine and levetiracetam. *Neurology* 69: 2076–2084.

Meador KJ, Gevins A, Leese PT, Otoul C and Loring DW (2010) Neurocognitive effects of Brivaracetam, Levetiracetam, and Lorazepam. *Epilepsia* 00: 0–0 (accessed 1 November 2010).

Mecarelli O, Piacenti A, Pulitano P, et al. (2001) Clinical and electroencephalographic effects of topiramate in patients with epilepsy and healthy controls. *Clin Neuropharmacol* 24: 284–289.

Mecarelli O, Vicenzini E, Pulitano P, et al. (2004) Clinical, cognitive, and neurophysiologic correlates of short-term treatment with carbamazepine, oxcarbazepine, and levetiracetam in healthy volunteers. *Ann Pharmacother* 38: 1816–1822.

Oken BS and Salinsky M (1992) Alertness and attention: Basic science and electrophysiologic correlates. *J Clin Neurophysiol* 9: 480–494.

Park MK, Belluzzi JD, Han SH, Cao J and Leslie FM (2007) Age, sex and early environment contribute to individual differences in nicotine/acetaldehyde-induced behavioral and endocrine responses in rats. *Pharmacol Biochem Behav* 86: 297–305.

Placidi F, Tombini M, Romigi A, et al. (2004) Topiramate: effect on EEG interictal abnormalities and background activity in patients affected by focal epilepsy. *Epilepsy Res* 58: 43–52.

Salinsky MC, Binder LM, Oken BS, Storzbach D, Aron CR and Dodrill CB (2002) Effects of gabapentin and carbamazepine on EEG and cognition in healthy volunteers. *Epilepsia* 43: 482–490.

Salinsky MC, Spencer DC, Oken BS and Storzbach D (2004) Effects of oxcarbazepine and phenytoin on the EEG and cognition in healthy volunteers. *Epilepsy Behav* 5: 894–902.

Sanquist TF, Rohrbaugh JW, Syndulko K and Lindsley DB (1980) Electrocortical signs of levels of processing: Perceptual analysis and recognition memory. *Psychophysiology* 17: 568–576.

Smith ME, Gevins A, Brown H, Karnik A and Du R (2001) Monitoring task load with multivariate EEG measures during complex forms of human computer interaction. *Hum Factors* 43: 366–380.

Smith ME, Gevins A, McEvoy LK, Meador KJ, Ray PG and Gilliam F (2006) Distinct cognitive neurophysiologic profiles for lamotrigine and topiramate. *Epilepsia* 47: 695–703.

Spreen O and Strauss E (1998) *A Compendium of Neuropsychological Tests: Administration, Norms, and Commentary*. New York: Oxford University Press.

Swanson JM and Volkow ND (2002) Pharmacokinetic and pharmacodynamic properties of stimulants: implications for the design of new treatments for ADHD. *Behav Brain Res* 130: 73–78.

Tou JT and Gonzalez RC (1974) *Pattern Recognition Principles*. Reading, MA: Addison-Wesley.

Van Reen E, Jenni OG and Carskadon MA (2006) Effects of alcohol on sleep and the sleep electroencephalogram in healthy young women. *Alcohol Clin Exp Res* 30: 974–981.