Pattern classification of working memory networks reveals differential effects of methylphenidate, atomoxetine and placebo in healthy volunteers

Andre Frank Marquand, Sara de Simoni, Owen G. O’Daly, Steven C. R. Williams, Janaina Mourao-Miranda, Mitul Mehta

To cite this version:
Andre Frank Marquand, Sara de Simoni, Owen G. O’Daly, Steven C. R. Williams, Janaina Mourao-Miranda, et al.. Pattern classification of working memory networks reveals differential effects of methylphenidate, atomoxetine and placebo in healthy volunteers. Neuropsychopharmacology, Nature Publishing Group, 2011, 10.1038/npp.2011.9. hal-00616527

HAL Id: hal-00616527
https://hal.archives-ouvertes.fr/hal-00616527
Submitted on 23 Aug 2011

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers. L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
Pattern classification of working memory networks reveals differential effects of methylenidate, atomoxetine and placebo in healthy volunteers

Andre F. Marquand¹*, MSc, Sara De Simoni¹, MSc, Owen G. O’Daly¹, PhD, Steven C. R. Williams¹, PhD, Janaina Mourão-Miranda²,¹, PhD and Mitul A. Mehta¹, PhD

¹: Department of Neuroimaging, Centre for Neuroimaging Sciences, Institute of Psychiatry, King’s College London.

2: Centre for Computational Statistics and Machine Learning, Department of Computer Science, University College London.

*: Corresponding Author. Andre Marquand, Department of Neuroimaging, Centre for Neuroimaging Sciences, Box P089, Institute of Psychiatry, De Crespigny Park, London, SE58AF, United Kingdom. Tel: (+44) 203 228 3066. Fax: (+44) 203 228 2116. Email: andre.marquand@kcl.ac.uk

Running Title: fMRI patterns for methylenidate and atomoxetine
Abstract

Stimulant and non-stimulant drugs can reduce symptoms of attention deficit/hyperactivity disorder (ADHD). The stimulant drug methylphenidate (MPH) and the non-stimulant drug atomoxetine (ATX) are both widely used for ADHD treatment, but their differential effects on human brain function remain unclear. We combined event-related fMRI with multivariate pattern recognition to characterise the effects of MPH and ATX in healthy volunteers performing a rewarded working memory (WM) task. The effects of MPH and ATX on WM were strongly dependent on their behavioural context. During non-rewarded trials only MPH could be discriminated from placebo (PLC), with MPH producing a similar activation pattern to reward. During rewarded trials both drugs produced the opposite effect to reward, i.e. attenuating WM networks and enhancing task related deactivations (TRDs) in regions consistent with the default mode network (DMN). The drugs could be directly discriminated during the delay component of rewarded trials: MPH produced greater activity in WM networks and ATX produced greater activity in the DMN. Our data provide evidence that: (1) MPH and ATX have prominent effects during rewarded WM in task activated and deactivated networks, (2) During the delay component of rewarded trials, MPH and ATX have opposing effects on activated and deactivated networks: MPH enhances TRDs more than ATX whereas ATX attenuates WM networks more than MPH and (3) MPH mimics reward during encoding. Thus, interactions between drug effects and motivational state are crucial in defining the effects of MPH and ATX.

Keywords: Methylphenidate, Atomoxetine, Working memory, Reward, Pattern recognition.
Introduction

Stimulant and non-stimulant medications that influence dopamine (DA) and noradrenaline (NA) neurotransmission can reduce symptoms of attention deficit/hyperactivity disorder (ADHD). The stimulant drug methylphenidate (MPH) has been shown to have consistently greater clinical efficacy than atomoxetine (ATX), a non-stimulant drug recently approved for the treatment of ADHD in the USA and Europe (Spencer et al., 1998; Michelson et al., 2001; Faraone et al., 2005; Kemner et al., 2005; Starr and Kemner, 2005; Newcorn et al., 2008). ATX nonetheless offers several potential advantages over MPH including reduced abuse liability, reduced risk of motor side effects and as an alternative treatment for patients non-responsive to stimulants (Newcorn et al., 2008). However, the mechanisms underlying their differences on human brain function are unclear.

There is converging evidence that weakened PFC function underlies several of the hallmark deficits in ADHD (Arnsten, 2006). In particular, working memory (WM) - the ability to hold and manipulate information for future action - is impaired in ADHD (Martinussen et al., 2005; Willcutt et al., 2005) and has been strongly linked to the activity of the catecholamines (DA and NA) within the PFC (Brozoski et al., 1979; Arnsten and Goldman-Rakic, 1985). WM performance is also known to be improved with MPH (Elliott et al., 1997; Bedard et al., 2004; Mehta et al., 2004), currently understood as resulting from an increased efficiency of fronto-parietal WM regions demonstrated using PET neuroimaging studies (Mehta et al., 2000; Schweitzer et al., 2004). Studies in experimental animals suggest ATX has a similar ability to improve WM function (Gamo et al., 2010), via effects on prefrontal cortical activity, although there are no comparative human neuroimaging studies of the effects of MPH and ATX on WM networks.

Previous studies in experimental animals have indicated that: (1) MPH inhibits both DA and NA transporters (DAT and NAT respectively; Seeman and Madras, 1998; Han and Gu, 2006), (2) ATX is a selective inhibitor of NAT (Wong et al., 1982; Bolden-Watson and Richelson, 1993) and (3) both
drugs increase concentrations of DA and NA in the prefrontal cortex (PFC), but only MPH increases DA in the striatum (Bymaster et al., 2002). However, the neural consequences of these differential actions in humans and their implications for functional brain networks are currently unknown.

Theoretically, systemically administered MPH and ATX may differentially influence distributed brain regions due to localised effects at DAT and NAT sites (Ciliax et al., 1999; Schou et al., 2005) and consequent effects on connected brain areas, in addition to the differential effects on striatal catecholamine neurotransmission demonstrated in rodents (Bymaster et al., 2002). Thus, differential effects of MPH and ATX may be distributed across multiple brain regions. Multivariate pattern recognition (PR) methods are sensitive to such spatially distributed information by making use of the correlation between brain voxels and afford substantially greater sensitivity than conventional mass-univariate analysis methods (Haynes and Rees, 2006; Kriegeskorte et al., 2006; Norman et al., 2006). Therefore, we combined event-related fMRI with a novel whole-brain PR analytic approach in order to characterise and discriminate acute effects of MPH and ATX in healthy volunteers performing a WM task. While we expected reductions in PFC activity after MPH, this study represents the first attempt to: (1) examine the effects of ATX on WM networks and (2) test potential differences between prefrontal cortical and striatal activation following administration of MPH and ATX in humans.

Finally, recent literature suggests an important contribution of reward to the regulation of WM-related brain activity (Ichihara-Takeda et al., 2010). This accords with evidence that both reward and MPH have similar effects on sustained attention task performance in ADHD (Trommer et al., 1991; Andreou et al., 2007). Therefore, we also explored the role of reward on WM function, with a focus on determining its impact on our ability to discriminate MPH and ATX.
Methods

Participant recruitment and data acquisition

Fifteen healthy male university students and members of the general public (aged 20-39) were recruited by local advertisement and were scanned on three occasions. Participants were screened by interview and physical exam for previous or current medical, psychiatric or neurological problems. Other exclusion criteria included any substance abuse history, smoking >5 cigarettes per day and consuming the equivalent of >5 cups of coffee per day. Participants were trained on the WM task on the screening day and were asked to refrain from alcohol and caffeine containing products for 24 hours prior to dosing. Participants provided written consent and the study was approved by South London Research Ethics Committee 3. On each scanning day, participants were screened for drugs of abuse and alcohol, then each participant received an oral dose of MPH (30mg), ATX (60mg), or a placebo (PLC) according to a randomised, double-blind Latin square design. Doses of MPH and ATX were chosen to approximately match doses used in clinical practice, and doses reported in the literature (e.g. Gilbert et al., 2006).

Scanning was performed on a on a General Electric Signa HDx 3T scanner and was timed to coincide with the peak plasma concentration for MPH and ATX (Wargin et al., 1983; Sauer et al., 2005). Between 90 and 135 minutes post-dose, six resting state arterial spin labelling scans were acquired, which will be reported separately. Approximately 135 minutes post-dose, gradient-echo (GE) echoplanar imaging was used to acquire 450 whole-brain images while participants performed a WM task (TR = 2s, TE = 30ms, FA = 75°, 38 3mm thick near-axial slices with 0.3mm gap, in-plane resolution = 3.75x3.75mm). A high-resolution GE structural scan was also acquired for each participant to assist accurate registration to a standard space (TR = 3s, TE = 30ms, FA = 90°, 43 3mm thick near axial slices with 0.3mm gap, in-plane resolution = 1.88x1.88mm).
Working memory task

During the WM task, 40 trials were presented with an inter-trial interval of 8 or 10 seconds and during each trial, participants were required to remember the spatial location of a target stimulus (a dot) relative to a fixation cross. The task allowed each WM component process (encoding, delay and retrieval) to be separately coded (Figure 1). Half the trials carried a monetary reward, indicated by the colour of the stimulus and the order of trials was randomised and counterbalanced across participants. During encoding (2 sec), the target stimulus was presented, followed immediately by a mask to disrupt visual iconic memory. After a variable length delay (7 or 9 sec), the target and an additional distractor stimulus were presented and participants indicated which of the stimuli matched the target location by left or right button press on a two-button response box (retrieval). At the conclusion of the trial, feedback was provided, indicating success or failure and responses and response time (RT) were recorded. Acquisition was optimised for volume-based PR, with stimuli presented in a TR-locked fashion, which ensures data vectors were sampled from approximately the same point on the haemodynamic response curve and helps to generate a consistent response pattern for each trial. The task was written in VB.net, presented via projector to a screen at the end of the scanner bed and viewed by participants through mirrors attached to the head coil. Participants completed a visual analogue scale (VAS; Bond and Lader, 1974) at four timepoints during each visit to record their subjective experience, which contained 16 items that were later collapsed to reflect two subjective factors: ‘alertness’ and ‘tranquillity’ (Herbert et al., 1976; supplementary material). Outside the scanner, VAS responses were measured with a ruler and inside the scanner a computerised VAS was administered where participants recorded their responses by moving a sliding cursor using the two-button response box.
FMRI data preprocessing

FMRI data were realigned, spatially normalised and smoothed with an isotropic 8mm Gaussian kernel using SPM5 (www.fil.ion.ucl.ac.uk/spm). Additional pre-processing was performed in Matlab (www.mathworks.com), which consisted of linearly detrending the data and applying a whole-brain mask to select intracerebral voxels. Classifier samples were constructed by: (1) shifting the onset of each trial by one volume to accommodate the haemodynamic delay, (2) converting brain volumes acquired during each task component to vectors and (3) averaging two (encoding, retrieval and shorter delay) or three (longer delay) consecutive volumes from each WM component. We averaged at least two volumes for each WM component to accommodate the temporal blurring induced by the haemodynamic response and to ensure that we captured the peak of the haemodynamic response. Trials where each participant responded incorrectly were excluded and remaining trials were averaged to construct a single mean sample per participant (averaged over approximately 16 correct trials). We constructed classifier samples for the baseline condition by extracting and averaging two volumes during the fixation period between trials (6-8 seconds after the end of feedback).

Classifier implementation

We used binary Gaussian process classifiers (GPCs; Rasmussen and Williams, 2006) to classify: (1) each WM component from baseline, (2) rewarded from non-rewarded trials and (3) each drug condition (ATX, MPH or PLC) from one another. GPCs are kernel classifiers similar to support vector machines (SVM) that have good performance for fMRI (Marquand et al., 2010b) and the main advantage of GPC over SVM is that GPCs provide probabilistic predictions and estimates of
predictive uncertainty. Theoretical background and implementation details for GPC have been presented elsewhere (Rasmussen and Williams, 2006; Marquand et al., 2010b) but a brief description is provided in supplementary material. In this work we use linear kernel GPCs that help prevent overfitting and allow direct extraction of the weight vector as an image.

Recursive Feature Elimination

We embedded all classifiers contrasting reward or drug state in a recursive feature elimination framework (RFE; Guyon et al., 2002), which is a backward elimination feature selection approach that aims to find a parsimonious set of features (voxels) by iteratively removing the least informative features. RFE was originally developed for SVM (SVM-RFE), and has been applied to multiple fMRI studies (e.g. De Martino et al., 2008; Formisano et al., 2008; Hanson and Halchenko, 2008) but here we adapt it to GPC ('GPC-RFE'; Marquand et al., 2010a). RFE starts by creating an ‘active feature set’, initially containing all cerebral voxels. A classifier is trained repeatedly on the active set and at each iteration features are ranked and a subset of the lowest ranking features is removed (2% of voxels), which continues until no features remain. Predictive performance is measured at each stage of feature removal on an independent sample, allowing an optimal number of features maximising predictive performance to be selected (supplementary material). RFE is most commonly applied because it modestly increases accuracy, but here our main motivation was because it yields a spatially sparse multivariate map (akin to a thresholded statistical parametric map), which is essential to prevent falsely inferring a brain region is functionally important when in fact it is not. RFE is a principled approach to achieve this goal and is more appropriate than an arbitrary voxel-wise threshold because it: (1) validates the multivariate pattern against predictive accuracy, (2) accommodate the multivariate structure of the pattern and (3) does not require specification of an arbitrary threshold level. We did not apply GPC-RFE to the classifiers contrasting task and baseline, because this is a trivial classification problem and the objective was only to define the brain activity...
pattern evoked by the task for which an unthresholded map is preferable, but for reference purposes, we provide classification accuracy for whole-brain classifiers trained to discriminate between all experimental contexts (supplementary material).

Cross-validation

RFE can be viewed as a model selection problem, where model complexity is determined by a single parameter (the number of features to retain) which must be set without using the test dataset to avoid overfitting. To achieve this, we used nested leave-one-out cross-validation (LOO-CV), which uses a three-way split of the data to provide an unbiased estimate of generalisation ability while also allowing unbiased parameter estimation. For each LOO-CV fold, we excluded all data for a single participant for the test set, then repeatedly repartitioned the remaining 14 participants into a validation set (1 participant) and training set (13 participants). We selected the optimal number of features on the validation set before applying it to the test set.

Visualisation of the differential activity pattern

To visualise the differential activity patterns, we retrained each GPC-RFE classifier using all participants’ data, where the optimal number of features was the mean across all training folds. For this application, we are interested in knowing how brain activity differs between experimental classes rather than providing a representation of the decision boundary, so we did not visualise classifier weights, which is common in PR (Mourao-Miranda et al., 2005). Instead, we employed a mapping approach that enables direct visualisation of the relative class distribution, where the coefficient scores at each voxel represent the relative difference between experimental classes in the context of the entire pattern (Marquand et al., 2010b; supplementary material).
Results

Performance measures

Repeated-measures ANOVA revealed that reaction time (RT) did not differ between drugs \( (F_{2,28} = 0.001, p = 0.99) \), or between rewarded and non-rewarded trials \( (F_{1,14} = 0.003, p = 0.96) \), and there was no reward \& drug interaction \( (F_{2,28} = 0.47, p = 0.63) \). Participants made fewer errors on rewarded relative to non-rewarded trials \( (F_{1,14} = 11.54, p < 0.01) \), but errors did not differ between drug conditions \( (F_{2,28} = 0.12, p = 0.89) \) and there was no reward \& drug interaction \( (F_{2,28} = 1.80, p = 0.19) \). A summary of RT and accuracy is provided in supplementary table S1.

Subjective measures

Several participants reported side effects to the administration of the drugs (e.g. nausea, drowsiness) but these were mild in all cases and mostly resolved prior to discharge on the study day. Subjective factor alertness and tranquillity were investigated as potential confounds to any drug effect using an independent repeated-measures ANOVA for each factor. For alertness, there was no main effect of drug \( (F_{2,28} = 0.59, p = 0.56) \), but a main effect of timepoint was observed \( (F_{1,14} = 8.83, p = 0.01) \), whereby post-dose VAS scores were slightly lower than pre-dose scores across all drug conditions. No drug \& timepoint interaction was found \( (F_{2,28} = 0.21, p = 0.81) \). For tranquillity, there was no main effect of drug \( (F_{2,28} = 1.78, p = 0.19) \) or timepoint \( (F_{1,14} = 0.02, p = 0.88) \) and no interaction effect \( (F_{2,28} = 0.48, p = 0.62) \).
Task networks

Whole-brain classifiers accurately discriminated each WM component process from baseline for all drug conditions (mean accuracy (SEM) of 18 classifiers: 97.61% (0.01); p < 0.01, binomial test). As noted, the magnitude of GPC coefficients at each voxel provides a measure of the relative difference in BOLD activation between classes in the context of the entire discriminating pattern and the sign indicates (‘favours’) the class with greater mean activation (Marquand et al., 2010b). GPC distribution maps (supplementary figures S1-2) revealed a distributed network (pattern) favouring the task component processes including bilateral intraparietal sulci (IPS; Brodmann area (BA) 7), middle frontal gyri (BA 9/46) and bilateral medial and inferior frontal gyri (BA 6 and 47 respectively) in addition to visual and motor cortical regions. The pattern favouring baseline (task related deactivations - TRDs) included regions comprising the default mode network (DMN), i.e. posterior cingulate cortex (PCC; BA30), precuneus (BA 31), medial prefrontal cortex (PFC; BA 9/10 and 32), and lateral parietal cortex (BA 39).

Classification of reward

Classification accuracy for GPC-RFE classifiers discriminating between rewarded and non-rewarded trials exceeded chance (50%) for all WM component processes and across all three drug conditions with the exception of the encoding component on MPH (mean (SEM) of six classifiers: 70.72% (0.04); Figure 2A). The pattern favoured reward and encompassed both the WM networks and TRDs described above. Specifically, BOLD activity in lateral PFC, parietal regions, medial PFC and PCC/precuneus was relatively increased (Figure 3; Figure S2); in other words, the effect of reward was to attenuate TRDs and enhance activity in the WM network. TRDs were most prominently attenuated during encoding and delay components of the rewarded WM task whereas visual and WM regions were most prominently enhanced during delay and retrieval components of the task
(See Figure 3). In summary, reward produces a generalised increase in BOLD activity, including both task-related activations (which increase with reward) and TRDs (which are suppressed with reward).

Classification accuracy for drug contrasts

For ATX vs. PLC, classification accuracy exceeded chance for encoding, delay and retrieval components of rewarded trials (p < 0.05), but not during any WM component for the non-rewarded trials (Figure 2B). For MPH vs. PLC, accuracy exceeded chance during encoding, delay and retrieval of rewarded trials and during encoding of non-rewarded trials (p < 0.05; Figure 2C). For MPH vs. ATX, classification accuracy exceeded chance for the delay component of rewarded trials (p < 0.05; Figure 2D).

For all classifiers exceeding chance, RT data were used to explore putative relationships between classifier performance and behaviour. No significant correlations between RT and GPC-RFE predictive probabilities were found. Note that correlations with accuracy were not appropriate because all participants were well trained and made only a small number of errors, and only correct trials were included in the image analysis.
Discriminating pattern for ATX vs. PLC (rewarded trials)

Maps derived from classifiers trained to discriminate ATX from PLC on rewarded trials (Figure 4) contained a distributed pattern favouring PLC that included WM networks and DMN; in other words, in the reward context, ATX attenuated BOLD activity in WM networks and enhanced TRDs. During encoding, the pattern favouring PLC included the DMN (medial PFC and PCC/precuneus) and WM networks (IPS and bilateral PFC - BA 9, 46 and 47). In addition, small clusters weakly favouring ATX were observed in the cerebellum and lateral PFC during encoding. During delay and retrieval components, the pattern favouring PLC was most prominent in WM regions.

Discriminating pattern for MPH vs. PLC (rewarded trials)

Maps derived from classifiers trained to discriminate MPH from PLC on rewarded trials (Figure 5) contained a distributed pattern favouring PLC similar to that observed for ATX, which also encompassed WM and DMN regions. During encoding the pattern favouring PLC was mostly localised to DMN regions, but during delay and retrieval, the PLC pattern additionally included clusters in WM, motor and visual regions and was most widespread during retrieval. The pattern favouring MPH was restricted to encoding and was localised mostly to the cerebellum and lateral PFC.
Discriminating pattern for MPH vs. PLC (non-rewarded trials)

The map derived from the classifier trained to discriminate MPH from PLC during the encoding component of non-rewarded trials (Figure 6) contained a distributed pattern, this time favouring MPH, including DMN, WM (e.g. IPS) and visual regions. Thus, in the absence of reward, MPH enhanced activity in WM networks and attenuated TRDs. Note that the map contrasting MPH and PLC shows a strong qualitative similarity to the one contrasting rewarded and non-rewarded trials in the encoding component of the PLC condition (Figure 3A).

[Figure 6 about here]

Discriminating pattern for MPH vs. ATX (rewarded trials)

The map derived from the classifier trained to discriminate MPH from ATX on the delay component of rewarded trials (Figure 7) contained distributed patterns favouring MPH and ATX. The pattern favouring MPH was mainly localised to WM regions (IPS and lateral PFC - BA 9/46) and the pattern favouring ATX was mainly localised to the DMN. Thus, during the delay component of rewarded trials MPH relative to ATX resulted in greater BOLD activity in WM networks, and ATX relative to MPH resulted in greater TRDs.

[Figure 7 about here]

For all contrasts, the differential patterns derived from GPC-RFE show a reasonably good correspondence to those derived from an equivalent univariate statistical parametric map (SPM),
except the SPM retained substantially fewer voxels (at p < 0.001, uncorrected for multiple comparisons) than were retained by the classifier (data not shown).

A concise summary of this complex set of results is provided in table 1.

[Table 1 about here]

**Discussion**

We have demonstrated differential effects of MPH and ATX on brain activity patterns in healthy volunteers performing a rewarded WM task. An important conclusion from our results is that the effects of MPH and ATX on WM are context-dependent. In the rewarded context, both MPH and ATX could be accurately discriminated from PLC across all task components, showing similar patterns of attenuation across the WM networks and enhanced TRDs. During the encoding component of non-rewarded trials, MPH, but not ATX, could be discriminated from PLC; MPH increased activity in WM regions and attenuated TRDs compared to PLC. The pattern of BOLD signal changes observed during the delay component of rewarded trials also discriminated MPH from ATX. In this context, and relative to ATX, MPH produced a pattern of increased activity in WM networks, while ATX produced greater activity in the DMN. Overall this complex set of findings suggests that: (1) both MPH and ATX have salient effects during rewarded WM in both task activated and deactivated networks, (2) During the delay component of rewarded trials, MPH and ATX had opposing effects on activated and deactivated networks and (3) MPH may mimic reward during encoding.

The results in this study were determined by applying recently developed PR techniques to the neuroimaging data, which afford substantially greater sensitivity than conventional mass-univariate
techniques (Haynes and Rees, 2006; Norman et al., 2006) by making use of spatial correlation between voxels, lending themselves well to whole-brain inference. These properties make PR ideally suited to drug discrimination studies, where drugs administered systemically can theoretically influence distributed brain regions due to direct effects at target receptor sites and consequent effects on connected brain regions. It is important to emphasize that multivariate brain maps derived from PR analysis provide a different perspective to mass-univariate analysis and should be interpreted differently. In particular, multivariate brain maps describe a pattern of activity, and coefficients should not be interpreted as representing focal effects because many brain regions potentially contribute to the accuracy of the classifier.

The WM networks identified in this study agree well with previous studies (Curtis et al., 2004; Gibbs and D’Esposito, 2005) and were sensitive to reward. During rewarded trials participants performed the task more accurately, which was reflected as a generalised pattern of increased brain activity throughout WM networks and in the DMN. Indeed, increased activity in WM brain regions is a known effect of reward on WM tasks (Pochon et al., 2002; Taylor et al., 2004; Pessoa and Engelmann, 2010) and may reflect an increase in neuronal effort.

Methylphenidate and atomoxetine did not alter performance accuracy or response latency during the WM task. However, previous studies using MPH and amphetamine have suggested that reductions in BOLD activation accompanied by equivalent behavioural performance reflect an increased efficiency of WM networks (Mattay et al., 2000; Mehta et al., 2000). Thus, for our data in a rewarded context, this would seem to be the most parsimonious explanation for the effects of MPH and ATX on task activation and deactivation networks. This effect is probably mediated by increased catecholamine concentrations in WM regions (Bymaster et al., 2002), which is known to focus neuronal activity by enhancing responses to task-relevant stimuli while suppressing background noise (Foote et al., 1975; Seamans et al., 2001). Historically, DA has been linked with WM performance through increasing the efficiency of PFC neurons by decreasing delay related response
to “noise” (Arnsten, 2007; Vijayraghavan et al., 2007) and the stabilisation of their sustained activity (Durstewitz et al., 2000). However, NA is probably also important as therapeutic doses of MPH increase PFC extracellular concentrations of NA substantially more than DA (Berridge et al., 2006), and the beneficial effects of MPH and ATX on WM can be blocked by either DA D1 or NA α2 receptor antagonists (Arnsten and Dudley, 2005; Gamo et al., 2010). NA is also known to increase delay-related activity of PFC neurons in response to “signals” (Arnsten, 2007) and increase the salience of novel stimuli leading to the suggestion that it serves as an alarm system for contextual changes (Yu and Dayan, 2005).

The PCC, precuneus and ventromedial PFC are known to show decreased activity during a wide range of goal-directed tasks (Shulman et al., 1997). These regions have been proposed to underlie a ‘default mode’ of brain function (Raichle et al., 2001) and it is thought that to facilitate goal-directed action task-irrelevant mental activity in these regions must be suppressed. Indeed, failure to suppress default mode activity reflects momentary lapses in attention (Weissman et al., 2006) resulting in increased probability of error (Eichele et al., 2008). There is also preliminary evidence that ADHD may be characterised by deficiencies in attentional focus and insufficient suppression of brain activity in focal regions of the DMN (Fassbender et al., 2009) and that MPH may normalise the amplitude of TRDs in treatment-responsive ADHD participants (Peterson et al., 2009). Our results are consistent with this interpretation and further show that suppression of task-irrelevant mental activity may be a mechanism common to both MPH and ATX. Importantly, this effect was context-dependent, as it was only observed during rewarded trials.

In a rewarded context, classification accuracy was equivalent for classifiers discriminating MPH or ATX from PLC for each WM component, although accuracy was slightly higher for both drugs during retrieval relative to encoding and delay. Qualitatively, the effects of MPH and ATX were comparable, with both drugs producing a generalised decrease in brain activity in WM networks and DMN (i.e. attenuation of activity in WM networks and enhancement of TRDs). Nonetheless, the
extent of these effects separated the drugs during the delay component of rewarded trials: ATX attenuated BOLD activity in WM networks more than MPH and MPH enhanced TRDs more than ATX.

Microdialysis studies in rodents have shown that MPH and ATX increase DA concentration in the PFC, but only MPH increases DA in the striatum (Bymaster et al., 2002) and that therapeutic doses of MPH increase catecholamine concentration in the PFC substantially more than in the striatum (Berridge et al., 2006). However, in our study we did not observe increased striatal activity following MPH, similar to other neuroimaging studies in healthy volunteers (Mehta et al., 2000; Udo de Haes et al., 2007). This may be because the WM task we employed does not substantially engage the striatum, even for rewarded trials (Supplementary Figure S2), which is consistent with a recent review of the effects of reward on WM (Pessoa and Engelmann, 2010) or simply because the consequential effects of MPH on striatal DA levels are expressed in connected brain regions. Thus, subcortical effects of MPH on DA remain a candidate mechanism for the differential effects of MPH and ATX, as the PFC and striatum are strongly connected by parallel corticostriatal circuits (Alexander et al., 1986) and there is emerging evidence suggesting that the striatal DA system plays a role in the modulation of the DMN (Kelly et al., 2009; Tomasi et al., 2009). However, studies concurrently measuring striatal DA release and its functional consequences on brain activity are required to test this hypothesis explicitly.

In a non-rewarded context, it was only possible to discriminate MPH from PLC during encoding. In this case, the differential pattern (Figure 6) bears a strong qualitative resemblance to that differentiating rewarded from non-rewarded trials (Figure 3) suggesting that while MPH didn’t improve performance at the dose administered MPH nevertheless mimics the reward effect. Discrimination accuracies for classifiers contrasting rewarded and non-rewarded trials were also consistently lower on MPH than on ATX or PLC and did not exceed chance for encoding, indicating that activity patterns discriminating reward and non-rewarded trials were less distinguishable on MPH (Figure 2A), which is consistent with the suggestion that MPH increases task salience (Volkow
et al., 2004). This effect is probably mediated by DA, because a learned association between a cue and a reward results in increased phasic dopaminergic firing during cue presentation not reward delivery (Schultz et al., 1993) and increased dopaminergic firing, often followed by immediate depression, is also associated with stimuli that resemble the rewarded stimulus (Schultz and Romo, 1990). Catecholaminergic signalling has also been associated with an ‘inverted-U’ dose-response relationship in the PFC (Arnsten, 2006; Levy, 2009) with optimal PFC function at intermediate concentrations and too much or too little DA or NA resulting in impaired PFC function. Although speculative at this stage, such a relationship may underlie different contextual effects of MPH, where rewarded and non-rewarded contexts may engage curves with different optimal dosing. Also, we only administered one dose of each drug here, so it is possible that ATX shares the reward emulating effect at a different dose, which could additionally account for the classifier’s inability to discriminate MPH and ATX during encoding of non-rewarded trials (Supplementary Figure S4).

Individual differences in response to drug administration may be an interesting line of future research. In particular, genetic factors influence responses to stimulants (Mattay et al., 2003) and although we did not collect genetic information here, inclusion of genetic factors can only be expected to improve predictive performance. As noted, only one dose of each drug was administered so dose effects cannot be excluded as confounds, but three lines of evidence speak against this possibility: first, administered doses were matched according to doses used in clinical practice. Second, motor evoked potentials were altered to a similar extent for both drugs using identical doses to those administered here (Gilbert et al., 2006). Third, opposing effects of MPH and ATX on activated and deactivated task networks during the delay component of rewarded trials are difficult to explain by a simple dose effect.

In summary, we accurately discriminated the effects of MPH and ATX on rewarded and non-rewarded WM networks using multivariate PR. We suggest that this method is ideal for drug discrimination studies because for most psychotropic medications subtle distributed effects
probably predominate over strong focal effects. More importantly, our results demonstrate that MPH and ATX have effects on WM function that are context-dependent and suggest that the interaction between drug effects and motivational state will be crucial in defining the beneficial effects of MPH and ATX in ADHD.

Disclosure/Conflict of Interest

The authors declare no conflict of interest.

Acknowledgements

The authors thank Bill Vennart, Paul Maguire and Caroline Woodlridge from Pfizer Global Research and Development and the Wellcome Trust and EPSRC for continued funding of the Centre for Neuroimaging Sciences. JMM was funded by the Wellcome Trust and AM gratefully acknowledges support from King’s College Annual Fund.
References

Alexander GE, Delong MR, Strick PL (1986) Parallel Organization of functionally segregated circuits linking basal ganglia and cortex. Annual Review of Neuroscience 9:357-381.

Andreou P, Neale BM, Chen W, Christiansen H, Gabriels I, Heise A, Meidad S, Muller UC, Uebel H, Banaschewski T, Manor I, Oades R, Roeyers H, Rothenberger A, Sham P, Steinhausen HC, Asherson P, Kuntsi J (2007) Reaction time performance in ADHD: improvement under fast-incentive condition and familial effects. Psychological Medicine 37:1703-1715.

Arnsten A, Dudley A (2005) Methylphenidate improves prefrontal cortical cognitive function through α2 adrenergoreceptors and dopamine D1 receptor actions: Relevance to therapeutic effects in Attention-Deficit Hyperactivity Disorder. Behavioural and Brain Functions 1:2-10.

Arnsten AFT (2006) Stimulants: Therapeutic actions in ADHD. Neuropsychopharmacology 31:2376-2383.

Arnsten AFT (2007) Catecholamine and second messenger Influences on prefrontal cortical networks of "representational knowledge": A rational bridge between genetics and the symptoms of mental illness. Cerebral Cortex 17:i6-i15.

Arnsten AFT, Goldman-Rakic PS (1985) Alpha-2-adrenergic mechanisms in prefrontal cortex associated with cognitive decline in aged non-human primates. Science 230:1273-1276.

Bedard AC, Martinussen R, Ickowicz A, Tannock R (2004) Methylphenidate improves visual-spatial memory in children with attention-deficit/hyperactivity disorder. Journal of the American Academy of Child and Adolescent Psychiatry 43:260-268.

Berridge CW, Devilbiss DM, Andrzezewski ME, Arnsten AFT, Kelley AE, Schmeichel B, Hamilton C, Spencer RC (2006) Methylphenidate preferentially increases catecholamine neurotransmission within the prefrontal cortex at low doses that enhance cognitive function. Biological Psychiatry 60:1111-1120.

Bolden-Watson C, Richelson E (1993) Blockade by newly-developed antidepressants of biogenic-amine uptake into rat-brain synaptosomes. Life Sciences 52:1023-1029.

Bond AJ, Lader MH (1974) The use of analogue scales in rating subjective feelings. British Journal of Medical Psychology 47:211-218.

Brozoski TJ, Brown RM, Rosvold HE, Goldman PS (1979) Cognitive deficit caused by regional depletion of dopamine in prefrontal cortex of rhesus-monkey. Science 205:929-932.

Bymaster FP, Katner JS, Nelson DL, Hemrick-Luecke SK, Threlkeld PG, Heiligenstein JH, Morin SM, Gehlert DR, Perry KW (2002) Atomoxetine increases extracellular levels of norepinephrine and dopamine in prefrontal cortex of rat: A potential mechanism for efficacy in Attention Deficit/Hyperactivity Disorder. Neuropsychopharmacology 27:699-711.

Ciliax BJ, Drash GW, Staley JK, Haber S, Mobley CJ, Miller GW, Mufson EJ, Mash DC, Levey AI (1999) Immunocytochemical localization of the dopamine transporter in human brain. Journal of Comparative Neurology 409:38-56.

Curtis CE, Rao VY, D’Esposito M (2004) Maintenance of spatial and motor codes during oculomotor delayed response tasks. Journal of Neuroscience 24:3944-3952.

De Martino F, Valente G, Staeren N, Ashburner J, Goebel R, Formisano E (2008) Combining multivariate voxel selection and support vector machines for mapping and classification of fMRI spatial patterns. Neuroimage 43:44-58.

Durstewitz D, Seamans JK, Sejnowski TJ (2000) Dopamine-mediated stabilization of delay-period activity in a network model of prefrontal cortex. Journal of Neurophysiology 83:1733-1750.

Eichele T, Debener S, Calhoun VD, Specht K, Engel AK, Hugdahl K, von Cramon DY, Ullsperger M (2008) Prediction of human errors by maladaptive changes in event-related brain networks. Proceedings of the National Academy of Sciences of the United States of America 105:6173-6178.
Elliott R, Sahakian BJ, Matthews K, Bannerjea A, Rimmer J, Robbins TW (1997) Effects of methylphenidate on spatial working memory and planning in healthy young adults. Psychopharmacology 131:196-206.

Faraone SV, Biederman J, Spencer T, Michelson D, Adler L, Reimherr F, Glatt SJ (2005) Efficacy of atomoxetine in adult attention-deficit/hyperactivity disorder: a drug-placebo response curve analysis. Behav Brain Funct 1:16.

Fassbender C, Zhang H, Buzy WM, Cortes CR, Mizuiri D, Beckett L, Schweitzer JB (2009) A lack of default network suppression is linked to increased distractibility in ADHD. Brain Research 1273:114-128.

Foote S, R F, AP O (1975) Effects of Putative neurotransmitters on neuronal activity in monkey auditory cortex. Brain Research 86:229-242.

Formisano E, De Martino F, Bonte M, Goebel R (2008) "Who" Is Saying "What"? Brain-Based Decoding of Human Voice and Speech. Science 322:970-973.

Gamo NJ, Wamg M, Arnsten AF (2010) methylphenidate and atomoxetine enhance prefrontal function through alpha-2 adrenergic and dopamine D1 receptors. Journal of the American Academy of Child and Adolescent Psychiatry 49:1011-1023.

Gibbs SEB, D'Esposito M (2005) A functional MRI study of the effects of bromocriptine, a dopamine receptor agonist, on component processes of working memory. Psychopharmacology 180:644-653.

Gilbert DL, Ridel KR, Sallee FR, Zhang J, Lipps TD, Wassermann EM (2006) Comparison of the inhibitory and excitatory effects of ADHD medications methylphenidate and atomoxetine on motor cortex. Neuropsychopharmacology 31:442-449.

Guyon I, Weston J, Barnhill S, Vapnik V (2002) Gene selection for cancer classification using support vector machines. Machine Learning 46:389-422.

Han DD, Gu HH (2006) Comparison of the monoamine transporters from human and mouse in their sensitivities to psychostimulant drugs. BMC Pharmacology 6:Article No.: 6.

Hanson SJ, Halchenko YO (2008) Brain reading using full brain support vector machines for object recognition: There is no "Face" identification area. Neural Computation 20:486-503.

Haynes JD, Rees G (2006) Decoding mental states from brain activity in humans. Nat Rev Neurosci 7:523-534.

Herbert M, Johns MW, Dore C (1976) Factor-analysis of analog scales measuring subjective feelings before and after sleep. British Journal of Medical Psychology 49:373-379.

Ichihara-Takeda S, Takeda K, Funahashi S (2010) Reward acts as a signal to control delay-period activity in delayed-response tasks. Neuroreport 21:367-370.

Kelly C, de Zubicaray G, Di Martino A, Copland DA, Reiss PT, Klein DF, Castellanos FX, Milham MP, McMahon K (2009) L-Dopa Modulates Functional Connectivity in Striatal Cognitive and Motor Networks: A Double-Blind Placebo-Controlled Study. Journal of Neuroscience 29:7364-7378.

Kemner JE, Starr HL, Ciccone PE, Hooper-Wood CG, Crockett RS (2005) Outcomes of OROS (R) methylphenidate compared with atomoxetine in children with ADHD: A multicenter, randomized prospective study. Advances in Therapy 22:498-512.

Kriegeskorte N, Goebel R, Bandettini P (2006) Information-based functional brain mapping. Proceedings of the National Academy of Sciences of the United States of America 103:3863-3868.

Levy F (2009) Dopamine vs noradrenaline: inverted-U effects and ADHD theories. Australian and New Zealand Journal of Psychiatry 43:101-108.

Marquand A, De Simoni S, O'Daly O, Mourao-Miranda J, Mehta M (2010a) Quantifying the information content of brain voxels using target information, Gaussian processes and recursive feature elimination Qu. In: International Conference on Pattern Recognition. Istanbul, Turkey.
Marquand A, Howard M, Brammer M, Chu C, Coen S, Mourao-Miranda J (2010b) Quantitative prediction of subjective pain intensity from whole-brain fMRI data using Gaussian processes. NeuroImage 49:2178-2189.

Martinussen R, Hayden J, Hogg-Johnson S, Tannock R (2005) A meta-analysis of working memory impairments in children with attention-deficit/hyperactivity disorder. Journal of the American Academy of Child and Adolescent Psychiatry 44:377-384.

Mattay VS, Callicott JH, Bertolino A, Heaton I, Frank JA, Coppola R, Berman KF, Goldberg TE, Weinberger DR (2000) Effects of dextroamphetamine on cognitive performance and cortical activation. NeuroImage 12:268-275.

Mattay VS, Goldberg TE, Fera F, Hariri AR, Tessitore A, Egan MF, Kolachana B, Callicott JH, Weinberger DR (2003) Catechol O-methyltransferase val(158)-met genotype and individual variation in the brain response to amphetamine. Proceedings of the National Academy of Sciences of the United States of America 100:6186-6191.

Mehta MA, Goodyer IM, Sahakian BJ (2004) Methylphenidate improves working memory and set-shifting in AD/HD: relationships to baseline memory capacity. Journal of Child Psychology and Psychiatry 45:293-305.

Mehta MA, Owen AM, Sahakian BJ, Mavaddat N, Pickard JD, Robbins TW (2000) Methylphenidate enhances working memory by modulating discrete frontal and parietal lobe regions in the human brain. Journal of Neuroscience 20:6.

Michelson D, Faries D, Wernicke J, Kelsey D, Kendrick K, Sallee FR, Spencer T, Atomoxetine ASG (2001) Atomoxetine in the treatment of children and adolescents with attention-deficit/hyperactivity disorder: A randomized, placebo-controlled, dose-response study. Pediatrics 108.

Mourao-Miranda J, Bokde AL, Born C, Hampel H, Stetter M (2005) Classifying brain states and determining the discriminating activation patterns: Support Vector Machine on functional MRI data. NeuroImage 28:980-995.

Newcorn JH, Kratochvil CJ, Allen AJ, Casat CD, Ruff DD, Moore RJ, Michelson D, Methylphenidate C (2008) Methylphenidate and osmotically released methylphenidate for the treatment of attention deficit hyperactivity disorder: Acute comparison and differential response. American Journal of Psychiatry 165:721-730.

Norman KA, Polyn SM, Detre GJ, Haxby JV (2006) Beyond mind-reading: multi-voxel pattern analysis of fMRI data. Trends Cogn Sci 10:424-430.

Pessoa L, Engelmann J (2010) Embedding reward signals into perception and cognition. Frontiers in Neuroscience 4:1-8.

Peterson BS, Potenza MN, Wang ZS, Zhu HT, Martin A, Marsh R, Plessen KJ, Yu S (2009) An fMRI Study of the Effects of Psychostimulants on Default-Mode Processing During Stroop Task Performance in Youths With ADHD. American Journal of Psychiatry 166:1286-1294.

Pochon JB, Levy R, Fossati P, Lehericy S, Poline JB, Pillon B, Le Bihan D, Dubois B (2002) The neural system that bridges reward and cognition in humans: An fMRI study. Proceedings of the National Academy of Sciences of the United States of America 99:5669-5674.

Raichle ME, MacLeod AM, Snyder AZ, Powers WJ, Gusnard DA, Shulman GL (2001) A default mode of brain function. Proceedings of the National Academy of Sciences of the United States of America 98:676-682.

Rasmussen C, Williams CKI (2006) Gaussian Processes for Machine Learning. Cambridge, Massachusetts: The MIT Press.

Sauer JM, Ring BJ, Witcher JW (2005) Clinical pharmacokinetics of atomoxetine. Clinical Pharmacokinetics 44:571-590.

Schou M, Hallidin C, Pike VW, Mozley PD, Dobson D, Innis RB, Farde L, Hall H (2005) Post-mortem human brain autoradiography of the norepinephrine transporter using (S,S)- F-18 FMeNER-D-2. European Neuropsychopharmacology 15:517-520.
Schultz W, Romo R (1990) Dopamine neurons of the monkey midbrain - contingencies of responses to stimuli eliciting immediate behavioural reactions. Journal of Neurophysiology 63:607-624.

Schultz W, Apicella P, Ljungberg T (1993) Responses of monkey dopamine neurons to reward and conditioned-stimuli during successive steps of learning a delayed-response task. Journal of Neuroscience 13:900-913.

Schweitzer JB, Lee DO, Hanford RB, Zink CF, Ely TD, Tagamets MA, Hoffman JM, Grafton ST, Kilts CD (2004) Effect of methylphenidate on executive functioning in adults with attention-deficit/hyperactivity disorder: Normalization of behavior but not related brain activity. Biological Psychiatry 56:597-606.

Seamans JK, Gorelova N, Durstewitz D, Yang CR (2001) Bidirectional dopamine modulation of GABAergic inhibition in prefrontal cortical pyramidal neurons. Journal of Neuroscience 21:3628-3638.

Seeman P, Madras BK (1998) Anti-hyperactivity medication: methylphenidate and amphetamine. Molecular Psychiatry 3:386-396.

Shulman GL, Fiez JA, Corbetta M, Buckner RL, Miezin FM, Raichle ME, Petersen SE (1997) Common blood flow changes across visual tasks. 2. Decreases in cerebral cortex. Journal of Cognitive Neuroscience 9:648-663.

Spencer T, Biederman J, Wilens T, Prince J, Hatch M, Jones J, Harding M, Faraone SV, Seidman L (1998) Effectiveness and tolerability of atomoxetine in adults with attention deficit hyperactivity disorder. American Journal of Psychiatry 155:693-695.

Starr HL, Kemner J (2005) Multicenter, randomized, open-label study of OROS methylphenidate versus atomoxetine: Treatment outcomes in African-American children with ADHD. Journal of the National Medical Association 97:115-165.

Taylor SF, Welsh RC, Wager TD, Phan KL, Fitzgerald KD, Gehring WJ (2004) A functional neuroimaging study of motivation and executive function. Neuroimage 21:1045-1054.

Tomasi D, Volkow ND, Wang RL, Telang F, Wang GJ, Chang L, Ernst T, Fowler JS (2009) Dopamine Transporters in Striatum Correlate with Deactivation in the Default Mode Network during Visuospatial Attention. Plos One 4.

Trommer BL, Hoeppner JAB, Zecker SG (1991) The go-no go test in attention-deficit disorder is sensitive to methylphenidate. Journal of Child Neurology 6:S128-S131.

Udo de Haes JJ, Maguire RP, Jager PL, Paans AMJ, den Boer JA (2007) Methylphenidate-induced activation of the anterior cingulate but not the striatum: a 15O H2O PET study in healthy volunteers. Hum Brain Mapp 28:625-635.

Vijayraghavan S, Wang M, Birnbaum SG, Williams GV, Arnsten AFT (2007) Inverted-U dopamine D1 receptor actions on prefrontal neurons engaged in working memory. Nature Neuroscience 10:376-384.

Volkow ND, Wang GJ, Fowler JS, Telang F, Maynard L, Logan J, Gatley SJ, Pappas N, Wong C, Vaska P, Zhu W, Swanson JM (2004) Evidence that methylphenidate enhances the saliency of a mathematical task by increasing dopamine in the human brain. American Journal of Psychiatry 161:1173-1180.

Wargin W, Patrick K, Kiltz C, Guaitieri CT, Ellington K, Mueller RA, Kraemer G, Breese GR (1983) PHARMACOKINETICS OF METHYLPHENIDATE IN MAN, RAT AND MONKEY. Journal of Pharmacology and Experimental Therapeutics 226:382-386.

Weissman DH, Roberts KC, Visscher KM, Woldorff MG (2006) The neural bases of momentary lapses in attention. Nature Neuroscience 9:971-978.

Willcutt EG, Doyle AE, Nigg JT, Faraone SV, Pennington BF (2005) Validity of the executive function theory of attention-deficit/hyperactivity disorder: A meta-analytic review. Biological Psychiatry 57:1336-1346.

Wong DT, Threlkeld PG, Best KL, Bymaster FP (1982) A new inhibitor of norepinephrine uptake devoid of affinity for receptors in rat-brain. Journal of Pharmacology and Experimental Therapeutics 222:61-65.
Yu AJ, Dayan P (2005) Uncertainty, neuromodulation, and attention. Neuron 46:681-692.
### Table 1: Summary of classification results

| Contrast                           | WM components exceeding chance                                                                 | Neuronal Effects                                                                 | Figure(s)          |
|------------------------------------|-------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------|--------------------|
| Reward vs. Control                 | Encoding, delay and retrieval for all drug contrasts except for encoding on MPH                 | Reward increased activity in WM networks and suppressed TRDs                     | Figure 3, Figure S2|
| ATX vs. PLC (rewarded trials)      | Encoding, delay and retrieval                                                                  | ATX decreased activity in WM networks and enhanced TRDs                         | Figure 4           |
| ATX vs. PLC (non-rewarded trials)  | None                                                                                           | -                                                                               | Figure S3          |
| MPH vs. PLC (rewarded trials)      | Encoding, delay and retrieval                                                                  | MPH decreased activity in WM networks and enhanced TRDs                         | Figure 5           |
| MPH vs. PLC (non-rewarded trials)  | Encoding only                                                                                  | MPH increased activity in WM networks and suppressed TRDs                       | Figure 6           |
| MPH vs. ATX (rewarded trials)      | Delay only                                                                                     | MPH had greater activity in WM regions relative to ATX and ATX resulted in greater TRDs | Figure 7           |
| MPH vs. ATX (non-rewarded trials)  | None                                                                                           | -                                                                               | -                  |
Figures

Figure 1: Delayed match to location WM task

Figure 2: Classification accuracy for GPC-RFE classifiers for A: rewarded vs. non-rewarded trials, B: ATX vs. PLC, C: MPH vs. PLC and D: MPH vs. ATX. Asterisks indicate results significantly different from chance, i.e. 50% (p < 0.05, binomial test).

Figure 3: GPC-RFE distribution maps for classifiers discriminating between rewarded and non-rewarded trials for each WM component (PLC arm). A: encoding B: delay C: retrieval. Maps were rescaled such that the absolute maximum coefficient score was +/-1. The magnitude of GPC coefficients provides a measure of the relative difference in BOLD activity between experimental classes (in the context of the entire pattern) and the sign favours the class with greater mean activity. A distributed pattern favouring reward can be observed that indicates: (1) reward increased activity throughout WM networks and across all WM component processes (2) reward attenuated TRDs in DMN regions, which was especially prominent during encoding and delay.

Figure 4: GPC-RFE distribution maps for classifiers discriminating between ATX and PLC conditions for each WM component (rewarded trials). A: encoding B: delay C: retrieval. A distributed pattern favouring PLC can be observed that indicates that in a rewarded context: (1) ATX attenuated activity throughout WM networks, which was most prominent during delay and retrieval (2) ATX enhanced TRDs in DMN regions across all WM component processes. The cerebellum was the only region favouring ATX and was only observed during encoding.
Figure 5: GPC-RFE distribution maps for classifiers discriminating between MPH and PLC for each WM component (rewarded trials). A: encoding B: delay C: retrieval. A distributed pattern favouring PLC can be observed that indicates that in a rewarded context: (1) MPH attenuated activity throughout WM networks, which was most prominent during delay and retrieval (2) MPH enhanced TRDs in DMN regions across all WM component processes. The only regions favouring MPH were found during encoding and included cerebellum and small regions of lateral PFC.

Figure 6: GPC-RFE distribution maps for classifiers discriminating between MPH and PLC for the encoding WM component (non-rewarded trials). A distributed pattern favouring MPH can be observed that indicates that during encoding and in a non-rewarded context MPH enhanced activity in some WM regions and enhanced TRDs in DMN regions.

Figure 7: GPC-RFE distribution maps for classifiers discriminating between MPH and ATX for the delay component of rewarded trials. Distributed patterns of activity favouring both MPH and ATX can be observed that indicate that in a rewarded context: (1) MPH enhanced activity throughout WM networks relative to ATX and (2) ATX enhanced TRDs in DMN regions.
Figure 1

| Fixation | Cue | Mask | Fixation | Response | Feedback | Fixation |
|----------|-----|------|----------|----------|----------|----------|
| 8 or 10 sec | 750 ms | 750 ms | 2 sec | 8 or 10 sec | |

Time
- Encoding
- Delay
- Retrieval

Haemodynamic delay

FMRI Volumes
- Encoding
- Delay
- Retrieval

Onsets shifted by one volume

Figure 2

A

**Rewarded vs. Non-rewarded**

| Accuracy | Encoding | Delay | Retrieval |
|----------|----------|-------|-----------|
| 0.80     | *        | *     | *         |
| 0.75     | *        | *     | *         |
| 0.70     | *        | *     | *         |
| 0.65     | *        | *     | *         |
| 0.60     | *        | *     | *         |
| 0.55     | *        | *     | *         |
| 0.50     | *        | *     | *         |
| 0.45     | *        | *     | *         |
| 0.40     | *        | *     | *         |
| 0.35     | *        | *     | *         |
| 0.30     |          |       |           |

B

**ATX vs. PLC**

| Accuracy | Encoding | Delay | Retrieval |
|----------|----------|-------|-----------|
| 0.80     | *        | *     | *         |
| 0.75     | *        | *     | *         |
| 0.70     | *        | *     | *         |
| 0.65     | *        | *     | *         |
| 0.60     | *        | *     | *         |
| 0.55     | *        | *     | *         |
| 0.50     | *        | *     | *         |
| 0.45     | *        | *     | *         |
| 0.40     | *        | *     | *         |
| 0.35     | *        | *     | *         |
| 0.30     |          |       |           |

C

**MPH vs. PLC**

| Accuracy | Encoding | Delay | Retrieval |
|----------|----------|-------|-----------|
| 0.80     | *        | *     | *         |
| 0.75     | *        | *     | *         |
| 0.70     | *        | *     | *         |
| 0.65     | *        | *     | *         |
| 0.60     | *        | *     | *         |
| 0.55     | *        | *     | *         |
| 0.50     | *        | *     | *         |
| 0.45     | *        | *     | *         |
| 0.40     | *        | *     | *         |
| 0.35     | *        | *     | *         |
| 0.30     |          |       |           |

D

**MPH vs. ATX**

| Accuracy | Encoding | Delay | Retrieval |
|----------|----------|-------|-----------|
| 0.80     | *        | *     | *         |
| 0.75     | *        | *     | *         |
| 0.70     | *        | *     | *         |
| 0.65     | *        | *     | *         |
| 0.60     | *        | *     | *         |
| 0.55     | *        | *     | *         |
| 0.50     | *        | *     | *         |
| 0.45     | *        | *     | *         |
| 0.40     | *        | *     | *         |
| 0.35     | *        | *     | *         |
| 0.30     |          |       |           |
## Tables

### Table 1: Summary of classification results

| Contrast                  | WM components exceeding chance                                                                 | Neuronal Effects                                                                 | Figure(s)            |
|---------------------------|-------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------|----------------------|
| Reward vs. Control        | Encoding, delay and retrieval for all drug contrasts except for encoding on MPH                | Reward increased activity in WM networks and suppressed TRDs                     | Figure 3, Figure S2  |
| ATX vs. PLC (rewarded trials) | Encoding, delay and retrieval                                                              | ATX decreased activity in WM networks and enhanced TRDs                        | Figure 4             |
| ATX vs. PLC (non-rewarded trials) | None                                                                                     |                                                                                  | Figure S3            |
| MPH vs. PLC (rewarded trials) | Encoding, delay and retrieval                                                              | MPH decreased activity in WM networks and enhanced TRDs                        | Figure 5             |
| MPH vs. PLC (non-rewarded trials) | Encoding only                                                                            | MPH increased activity in WM networks and suppressed TRDs                     | Figure 6             |
| MPH vs. ATX (rewarded trials) | Delay only                                                                               | MPH had greater activity in WM regions relative to ATX and ATX resulted in greater TRDs | Figure 7             |
| MPH vs. ATX (non-rewarded trials) | None                                                                                     |                                                                                  | -                    |
