STABILITY OF FMRI STRIATAL RESPONSE TO ALCOHOL CUES: A HIERARCHICAL LINEAR MODELING APPROACH

Joseph P. Schacht, Ph.D.[1], **Raymond F. Anton, M.D.**[1], **Patrick K. Randall, Ph.D.**[1], **Xingbao Li, M.D.**[1,2], **Scott Henderson, B.A.**[1], and **Hugh Myrick, M.D.**[1,3]

[1]Charleston Alcohol Research Center, Department of Psychiatry and Behavioral Sciences, Medical University of South Carolina

[2]Center for Advanced Imaging Research, Medical University of South Carolina

[3]Ralph H. Johnson Department of Veterans Affairs Medical Center, Charleston, SC, USA

Abstract

In functional magnetic resonance imaging (fMRI) studies of alcohol-dependent individuals, alcohol cues elicit activation of the ventral and dorsal aspects of the striatum (VS and DS), which are believed to underlie aspects of reward learning critical to the initiation and maintenance of alcohol dependence. Cue-elicited striatal activation may represent a biological substrate through which treatment efficacy may be measured. However, to be useful for this purpose, VS or DS activation must first demonstrate stability across time. Using hierarchical linear modeling (HLM), this study tested the stability of cue-elicited activation in anatomically and functionally defined regions of interest in bilateral VS and DS. Nine non-treatment-seeking alcohol-dependent participants twice completed an alcohol cue reactivity task during two fMRI scans separated by 14 days. HLM analyses demonstrated that, across all participants, alcohol cues elicited significant activation in each of the regions of interest. At the group level, these activations attenuated slightly between scans, but session-wise differences were not significant. Within-participants stability was best in the anatomically defined right VS and DS and in a functionally defined region that encompassed right caudate and putamen (intraclass correlation coefficients of .75, .81, and .76, respectively). Thus, within this small sample, alcohol cue-elicited fMRI activation had good reliability in the right striatum, though a larger sample is necessary to ensure generalizability and further evaluate stability. This study also demonstrates the utility of HLM analytic techniques for serial fMRI studies, in which separating within-participants variance (individual changes in activation) from between-participants factors (time or treatment) is critical.

Keywords

alcohol; fMRI; cue reactivity; ventral striatum; dorsal striatum; HLM

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

1. Introduction

Alcohol cue reactivity (ACR), elicited among alcohol-dependent individuals upon exposure to alcohol-related stimuli (e.g., pictures, odors, tastes), has long been posited to reflect neuroadaptation engendered by classically and instrumentally conditioned associations between these stimuli and the hedonic effects of alcohol (e.g., Monti et al., 1987). The development of functional neuroimaging paradigms to assess ACR has yielded 20 published studies of this phenomenon, the majority of which have used functional magnetic resonance imaging (fMRI) to elucidate the brain areas activated by exposure to alcohol cues. Areas that consistently demonstrate such activation across different types of participants, sensory modalities, and fMRI paradigms include the anterior cingulate, medial and dorsolateral prefrontal and orbitofrontal cortices and the striatum (consisting of the nucleus accumbens [NAc], caudate, putamen, and pallidum) (for reviews, see Tapert et al., 2004; Sinha & Li, 2007). Of these areas, regions of the striatum have held particular interest in the alcohol and addictions fields, as they are believed to directly mediate either the learning mechanisms that allow the brain to associate the experience of reward with the stimuli that temporally precede it (Kelley, 2004; O'Doherty et al., 2004) or the attribution of incentive salience ("wanting") to such stimuli (Berridge & Robinson, 1998).

The striatum is anatomically and functionally divided into ventral (VS) and dorsal (DS) aspects; the former encompasses NAc and ventral caudate, and the latter dorsal caudate and putamen. Both are primarily innervated by midbrain dopaminergic (DA) nuclei: VS by the ventral tegmental area along the mesolimbic pathway, and DS by the substantia nigra along the nigrostriatal pathway. In rats, alcohol consumption induces DA release in both VS (particularly in the NAc; Imperato & Di Chiara, 1986; Yoshimoto et al., 1992) and DS (Melendez, Rodd-Henricks, McBride, & Murphy, 2003), though this release may have differential impact in each area. VS has relatively low tonic DA, and is thus more responsive to phasic DA stimulation (Zhang et al., 2009), an effect thought to underlie encoding of the reward value of salient, unpredicted events (Schultz, Dayan, & Montague, 1997). Such encoding may be particularly relevant to the initiation of alcohol use, when stimuli coincident with alcohol use have not yet been conditioned to its effects. In contrast, phasic DA stimulation of DS, where tonic DA is high, may contribute more to habit learning and maintenance (Balleine & O'Doherty, 2010), and ultimately to the automatic, compulsive alcohol seeking characteristic of end-stage alcohol dependence (Tiffany, 1990; Volkow et al., 2006). Thus, with respect to reward learning, VS is believed to underlie reward prediction (or stimulus-reward learning), while DS is believed to mediate the undertaking of action to achieve reward (or stimulus-response-reward learning) (Atallah et al., 2007; O'Doherty et al., 2004).

In fMRI studies of ACR, cue-elicited activations in both the midbrain DA nuclei (Filbey et al., 2008; Kareken et al., 2004; Myrick et al., 2004, 2010) and VS and DS (Braus et al., 2001; Filbey et al., 2008; Grüsser et al., 2004; Heinz et al., 2004; Kareken et al., 2004; Myrick et al., 2004, [2008], 2010; Völlstadt-Klein et al., 2010; Wrase et al., 2002, 2007) have frequently been reported, and are generally interpreted as conditioned anticipation of alcohol reward. In support of the putative VS-DS dissociation between salience and habit-learning, alcohol cue-elicited VS activation has been correlated with the magnitude of *in vivo* craving (Myrick et al., 2004, 2008), while DS activation has been correlated with the severity of alcohol dependence (Filbey et al., 2008). Further, heavy drinkers display greater cue-induced DS activation but less VS activation than light social drinkers (Völlstadt-Klein et al., 2010); among heavy drinkers, obsessive-compulsive craving for alcohol, a construct associated with more severe alcohol dependence (Anton, Moak, & Latham, 1996), correlates negatively with cue-elicited activation of VS, but positively with cue-elicited activation of DS (Völlstadt-Klein et al., 2010).

Some investigators have suggested that cue-elicited striatal activation may reflect not only the severity of craving or dependence, but also treatment efficacy or relapse potential (e.g., Mann et al., 2009). Several have studied the effects of prospective pharmacotherapies for alcohol dependence on such activation with placebo-controlled trials in which patients are scanned after a period of treatment (George et al., 2008; Grüsser et al., 2004; Hermann et al., 2006; Myrick et al., 2008, 2010). Relative to placebo, VS activation is significantly reduced when alcohol-dependent individuals are given the mu opioid partial antagonist naltrexone (Myrick et al., 2008) or the DA partial agonist aripiprazole (Myrick et al., 2010). Further, individuals who demonstrate greater cue-elicited VS activation after alcohol detoxification are more likely to relapse to heavy drinking (Grüsser et al., 2004). However, the most stringent test of the utility of cue-elicited fMRI activation for evaluation of treatment efficacy or relapse potential would require that individuals be scanned twice, at treatment initiation and again at its conclusion, and would assume that, absent treatment, cue-elicited activation is stable across multiple scan sessions (i.e., that there is no change in activation as a function of cue habituation, task learning, or noise inherent to the fMRI signal). Yet, no studies of the test-retest reliability (stability) of fMRI ACR paradigms exist.

The stability of fMRI activation has been assessed for a variety of other paradigms, from simple motor tasks (Havel et al., 2006) to more complex designs intended to model cognitive constructs such as anticipatory anxiety (Schunck et al., 2008) or episodic memory (Clément & Belleville, 2009). A number of statistical approaches have been proposed to measure fMRI test-retest reliability (see, e.g., Genovese et al., 1997; Maitra et al., 2002; Noll et al., 1997; Rombouts et al., 1998), but many have focused on the stability of whole brain (rather than region of interest, or ROI) results, and nearly all have used statistically thresholded images, which can arbitrarily exclude marginally subthreshold voxels from analysis and artificially truncate between-participants differences (though some studies have analyzed both anatomically- and statistically-defined ROIs; see Johnston et al., 2005; Manuck et al., 2007). One approach that is particularly amenable to testing the reliability of ROI activations from fMRI block designs is hierarchical linear modeling (HLM; Raudenbush & Bryk, 2002), which we have previously employed in other fMRI analyses (Myrick et al., 2008, 2010) and which has also been used in longitudinal analyses of structural MRI data (Frost, Kenward, & Fox, 2004; Yeh et al., 2007). HLM is useful for data structures in which measurements are "nested" within larger "bins" (e.g., time points within blocks, blocks within participants or scans within participants). ROI timecourse data, which model the blood-oxygen-level-dependent (BOLD) response in a particular brain area throughout a scan, are inherently such a structure; hundreds of time points are nested within each participant. HLM differs from traditional repeated measures analysis of variance (ANOVA) in that differences between blocks of observations made under different conditions (different stimuli) can be modeled as random coefficients, so that each participant has a separate coefficient. This greatly relaxes some of the restrictive assumptions of repeated measures ANOVA. Thus, HLM provides better estimates of both within-participants activations and the stability of these activations by properly modeling variance that is inherent to participants and to time. Further, through use of timecourse data, HLM can consider a fuller account of ROI activation than is possible with thresholded images.

This study aimed to test the stability of alcohol cue-elicited fMRI striatal activation using HLM. As our group and others have previously demonstrated such activation in VS and DS among non-treatment-seeking alcohol-dependent individuals, we hypothesized that visual alcohol cues would again elicit greater activation than control cues in anatomically and functionally defined regions of interest encompassing bilateral VS/DS. Further, we posited that the magnitude of cue-elicited activation in these ROIs would not change significantly between two scans separated by two weeks, and that within-participants activations would have good test-retest reliability.

2. Material and methods

2.1 Participants and procedure

All procedures were approved by the Institutional Review Board for Human Research at the Medical University of South Carolina. Ten non-treatment-seeking alcohol-dependent participants (6 males, 4 females) were recruited from the community using print advertisements. All participants met *Diagnostic and Statistical Manual of Mental Disorders, 4th edition* (DSM-IV) diagnostic criteria for Alcohol Dependence, including loss of control over drinking or an inability to cut down or quit, but they denied any active involvement in, or desire for, alcohol treatment. Exclusion criteria for all participants were: current DSM-IV abuse or dependence for other substances, other current Axis I disorders, psychoactive medication or substance use (except marijuana and nicotine) in the past 30 days by self-report or a positive urine drug screen, current suicidal or homicidal ideation, history of alcohol-related medical illness, or significant health problems. All participants were screened for DSM-IV criteria using the *Structured Clinical Interview for DSM-IV*. One female participant did not return for the second session, so her first scan data were not used (see Table 1 for demographic data for the final sample).

Participants were initially screened over the phone to determine whether they met inclusion criteria, and eligible participants then attended two laboratory sessions. They were told to abstain from drinking for 24 hours before both sessions. During the first session, handedness was assessed using the Edinburgh Handedness Inventory (Oldfield, 1971). One participant was left-handed, and to account for the influence of handedness, all analyses described below were also conducted with this participant excluded; there were no significant differences in any results. At the first session, participants completed two measures that assess alcohol dependence severity, the Alcohol Dependence Scale (ADS; Skinner & Allen, 1982) and Obsessive-Compulsive Drinking Scale (OCDS; Anton et al., 1996), and one that assesses craving for alcohol, the Alcohol Urge Questionnaire (AUQ; Bohn et al., 1995). Their past-month drinking was assessed using the Timeline Follow-Back (TLFB; Sobell & Sobell, 1992), a calendar-assisted, researcher-administered interview. To assess alcohol withdrawal (AW), the Clinical Institute Withdrawal Assessment for Alcohol-Revised (CIWA-Ar; Sullivan et al., 1989) was also administered. Participants were breathalyzed, and a rapid urine drug screen was obtained. No participant had evidence of alcohol use, a positive urine drug screen, or AW symptoms before either scan. Participants were then administered a cue-elicited craving task (from Myrick et al., 2004, 2008, 2010; see below) while undergoing a BOLD fMRI scan. After the scan, they were discharged home, without any instructions regarding their alcohol use except a reminder to abstain for 24 hours before the next session. Participants returned to the laboratory exactly 14 days later and were again breathalyzed, drug screened, and completed the AUQ and CIWA-Ar. Their drinking over the past 14 days was assessed using the TLFB and they were again scanned while undergoing the craving task before being compensated and discharged home.

2.2 Task

At each session, participants were positioned supine in the fMRI scanner and were administered a sip (10 mL) of their preferred alcohol in non-carbonated juice immediately before the scan began. During the scan, participants were shown interspersed visual stimuli that included images of alcoholic and non-alcoholic ("neutral") beverages, blurred versions of both of these kinds of images that lack object recognition, and a fixation cross. These images were selected from the Normative Appetitive Picture System ($n = 38$), supplemented with 22 additional images selected from print advertisements, and matched by color and hue. Stimuli were presented in six 120-s epochs (one 12 m-long run). Each epoch consisted of four 24-s blocks (one block each of alcohol, neutral beverage and blurred images, and

fixation). Each 24-s block was composed of five individual pictures, each displayed for approximately 4.8 s. Each block was followed by a 6 s washout period, which was intended to allow the hemodynamic response from the previous stimulus block to decline before the next block was presented. The alcohol images were specific to a beverage type (i.e., beer, wine, or liquor), with two blocks of each type. During the washout periods, participants were presented with a visual prompt that asked them to assess their "urge to consume alcohol"; data related to this assessment are not presented here.

2.3 Image acquisition and analysis

All functional images were acquired with a 3T Siemens TIM Trio scanner, using a gradient echo, echo-planar imaging (EPI) scan sequence. Image acquisition parameters were: repetition time = 2200 ms; echo time = 35 ms; flip angle = 90°; field of view = 192 mm; 36 slices 3.0 mm-thick, no gap. This sequence yielded 328 volumes for each 12-m run. Images were preprocessed using Statistical Parametric Mapping 5 software (SPM5; www.fil.ion.ucl.ac.uk/spm), implemented in MATLAB 7.6. First, images were realigned to the first functional volume, after which all participants had less than 2 mm of translational and 2° of rotational movement. Images were then stereotactically normalized to the Montreal Neurological Institute (MNI) average EPI template, resampled to 3 mm^3 voxel size, and spatially smoothed with an 8 mm^3 anisotropic Gaussian kernel. Using the AR(1) function in SPM5, images were also spatially whitened to account for correlations between nearby voxels. To eliminate low-frequency noise in the BOLD signal, a high-pass filter with a period of 240 s (i.e., twice the epoch length, to eliminate block-related drift) was used. A boxcar function that represented stimulus presentation and duration times was then convolved with the canonical hemodynamic response function to create a basis function for general linear modeling (GLM). To eliminate task-correlated motion, for each participant, the six motion parameters from the realignment were also included in this first-level model.

2.3.1 Anatomically defined regions of interest—Four anatomical ROIs, representing right and left VS and DS, were defined. Each ROI was a sphere with 6 mm radius, comprising thirty-three 3 mm^3 voxels, centered on the following MNI coordinates: right VS, [12, 15, −6]; left VS, [−12, 15, −6]; right DS, [12, 15, 6]; and left DS, [−12, 15, 6] (see Figure 1). These center coordinates were selected using the Harvard-Oxford Subcortical Structural Atlas (implemented in FSLView 3.1.8; see http://www.cma.mgh.harvard.edu/fsl_atlas.html), a probabilistic atlas of MNI space based on the brains of 37 individuals segmented into 21 subcortical structural regions, affine-transformed into MNI space, and then combined to create population probability maps that estimate the likelihood that a given coordinate represents a specific structure. The right and left VS center coordinates are within 1 mm of the points with the highest Harvard-Oxford atlas likelihoods of falling within the right and left NAcc (80% and 86% respectively) and the right and left DS coordinates are within 1 mm of the points with the highest likelihood of falling within the right and left anterior caudate (99% and 97% respectively). Each probability estimate is the highest in the atlas for each structure.

2.3.2 Functionally defined regions of interest—To complement the anatomically defined ROIs, functional ROIs within the striatum were also calculated. After pre-processing, participants' data were analyzed using the GLM described above, which yielded a first-level contrast image (alcohol vs. neutral beverage) for each participant. The contrast images from each scan were then combined and masked with a binary image of the striatum, derived from the Automated Anatomical Labeling library (Tzourio-Mazoyer et al., 2002). For each set of scans, a second-level random effects analysis was used to identify clusters within the striatum in which the alcohol vs. neutral beverage contrast was significantly nonzero. Group images were thresholded using a voxel-wise height threshold of $z > 1.96$, a

cluster probability of $p < 0.05$, corrected for multiple comparisons, and a cluster extent threshold (k) of 25 voxels. To determine the stability of these functionally defined ROIs, the clusters identified in Scan 1 were then used to mask participants' Scan 2 alcohol vs. neutral beverage contrast images. To identify areas of difference within the striatum between the two scans, a paired-samples t-test of the contrast images from each scan was also performed using the same cluster thresholding.

2.3.3 HLM analysis—HLM analysis was conducted for both kinds of ROIs. For the anatomically defined regions, using SPM5, the unadjusted first eigenvariate of the BOLD signal timecourse was extracted from each ROI in the first-level images. The first eigenvariate resulted from a singular value decomposition analysis of the first-level modeled BOLD signal across all 33 voxels in each ROI. It is the value that accounts for the greatest amount of variance in the BOLD signal across these voxels, accounting for their spatial relation to each other (see Friston et al., 2006). For the functionally defined regions, the unadjusted first eigenvariate of the BOLD timecouse was extracted from the significant clusters in the first-level Scan 1 images and from the same regions (i.e., from the mask of the Scan 1 clusters) in the first-level Scan 2 images.

Using the statistical software package HLM 6.0.8 (Raudenbush, Bryk, & Congdon, 2009), the Scan 1 and Scan 2 timecourses from each ROI were then entered as the outcome variable in a series of three-level HLMs, with stimulus (Level 1) nested within session (Level 2: Scan 1 vs. 2) nested within participant (Level 3). One HLM was constructed for each ROI. The Level 1 parameters (alcohol, beverage, blur, and fixation), which modeled each participant's individual timecourse, were dummy-coded. By excluding any one of them from the model, it was thus possible to assess the difference in the outcome variable (BOLD activation in the ROI) between the excluded parameter (i.e., beverage) and any of the included parameters (i.e., alcohol). Both the Level 1 (stimulus-induced activation: alcohol vs. neutral beverage) and Level 2 (session) effects of interest were analyzed as random contrasts (i.e., their slopes were allowed to vary between participants). Given the small, *a priori* anatomically defined ROIs, an α threshold of $p < 0.05$ was used to determine the statistical significance of stimulus-induced activation and between-session reliability for these regions.

Individual parameter estimates (i.e., slopes) for the alcohol vs. neutral beverage contrast were derived by adding each participant's Bayesian residual estimate from the HLM to the group parameter estimate for the slope in each scan. To determine the reliability of these individual parameter estimates between sessions, the intraclass correlation coefficient (ICC), which tests the consistency of individual measurements by relating between- and within-participant variance, was calculated. Specifically, the two-way mixed single measures coefficient, $ICC_{(3, 1)}$ (Shrout & Fleiss, 1979), which treats measurement time points as fixed factors and the measurements themselves as random factors, was computed. This measure has been established as a sufficient test of functional imaging test-retest reliability (Aron et al., 2006; Caceres et al., 2009; Fliessbach et al., 2010), and allows generalization beyond the study population.

3. Results

3.1 Behavioral data

There were no significant differences in participants' drinking during the month before the first scan (baseline) and during the 14 days between scans (paired samples t-test; $t(8) = -0.064$, $p = 0.95$ for drinks per drinking day; $t(8) = 0.394$, $p = 0.71$ for percent days abstinent; see Table 1). There were also no significant differences in participants' craving for alcohol (AUQ score) before each scan (paired samples t-test; $t(8) = -0.81$, $p = 0.44$).

3.2 Anatomically defined ROIs

The HLM analysis indicated that, as compared to neutral beverage pictures, alcohol pictures elicited significantly greater average activation in all of the anatomically defined ROIs in both Scan 1 and 2, though the magnitude of this activation was slightly smaller in all ROIs in Scan 2. Table 2 displays, for each ROI, t statistics that reflect the magnitude of the groupwise HLM-derived alcohol vs. neutral beverage contrast for each session, as well as the magnitude of the difference in this contrast between Scans 1 and 2. There was no significant difference between sessions in the magnitude of group-wise activation in any of the ROIs (i.e., none of the Level 2 session contrasts were significant). The HLM-derived individual alcohol vs. neutral beverage parameter estimates for each ROI had varying degrees of within-participants stability (see Table 2 for ICCs for each ROI; the p values that accompany these figures refer to the significance of their difference from zero). Figure 2 displays these individual parameter estimates graphically for each ROI. Of note, one participant displayed little activation in right VS and DS; exclusion of this participant from the analysis reduced within-participants reliability significantly in right VS (from $ICC_{(3, 1)}$ = .752, p = .006 to $ICC_{(3, 1)}$ = .501, p = .09), but not in right DS.

3.3 Functionally defined ROIs

In Scan 1, across the entire striatum, there were two significant clusters of alcohol vs. neutral beverage activation: one, comprising 136 voxels, that encompassed the left anterior caudate and putamen (MNI coordinates of the maximum voxel = [−18, 15, 12], z = 3.66, corrected cluster-level p = .005); and one, comprising 71 voxels, that encompassed the right anterior caudate and putamen (maximum voxel = [12, 9, −3], z = 3.84, corrected cluster-level p = .05). In Scan 2, there was one significant cluster of activation, comprising 141 voxels, that encompassed the right caudate and putamen (maximum voxel = [18, 12, 12], z = 3.91, corrected cluster-level p = .01) and overlapped with the Scan 1 right-sided cluster; no cluster in the left striatum survived thresholding in Scan 2 (see Figure 3). The right-sided Scan 2 cluster extended further in the dorsal and posterior directions than the Scan 1 cluster. The paired-samples t-test of the Scan 1 vs. Scan 2 contrast images identified no areas in which there was greater activation in Scan 1 than Scan 2, and one cluster, comprising 77 voxels in the posterior dorsal putamen, in which there was greater activation in Scan 2 than Scan 1 (maximum voxel = [27, −9, 0], z = 4.08, corrected cluster-level p = .05). Uncorrected cluster-level statistics for all clusters were significant at $p < .005$.

The HLM analysis of the two significant clusters identified in Scan 1 and the mask of these clusters in Scan 2 indicated significant alcohol vs. neutral beverage activation in both clusters in both sessions, though activation was again slightly attenuated in Scan 2. For these clusters, Table 2 displays t statistics that reflect the magnitude of the groupwise HLM-derived alcohol vs. neutral beverage contrast, and the magnitude of the difference in this contrast between Scans 1 and 2. Consistent with the anatomically defined ROIs, there was no significant difference in the magnitude of activation in either cluster between sessions. While the HLM-derived group parameter estimate for activation in the Scan 1-defined left-sided cluster was significant in Scan 2 (p = .013), there were not enough contiguous suprathreshold voxels in the left striatum to survive cluster thresholding in the functional analysis of the Scan 2 images. The ICCs for the HLM-derived individual parameter estimates for each functionally defined ROI were similar to those for the anatomically defined ROIs, and are also displayed in Table 2 and Figure 2. Exclusion of the participant who displayed little activation in the right-sided functional cluster did not significantly reduce the reliability estimate for this cluster.

4. Discussion

These data indicate that, in this small sample, (1) alcohol cues elicited significant activation, relative to neutral beverage cues, within subregions of the striatum; (2) at the group level, these activations attenuated slightly, but remained stable, across a period of two weeks; and (3) specific ROIs, specifically the anatomically defined right ventral and dorsal striatum and a functionally defined cluster encompassing both areas, exhibited good within-participants reliability. VS and DS are believed to mediate aspects of reward learning critical to the initiation and maintenance of alcohol dependence and other addictive behaviors. If within-participants measurements of their activation are reliable, statistical power to detect a between-scans treatment effect would be significantly increased, and the number of participants necessary for an adequately-powered treatment outcome study would be reduced (e.g., Potvin & Schultz, 2000).

Though many functional imaging studies of alcohol dependence and other addictive behaviors have highlighted the importance of striatal reward-related activation (e.g., Beck et al., 2009; Braus et al., 2001; Hariri et al., 2006; Heinz et al., 2004; Park et al., 2010; Völlstadt-Klein et al., 2010; Wrase et al., 2007), few have attempted to assess the stability of this construct. Importantly, this study employed anatomically defined ROIs, and centered them in brain regions with clear biological relevance for alcohol dependence (rather than in areas activated in a whole brain analysis that likely reflect other aspects of reward-related tasks, such as visual attention or cognitive appraisal). Fliessbach and colleagues (2010) recently reported poor to fair reliability for fMRI activation in similarly anatomically-defined ventral striatal and orbitofrontal cortex ROIs during reward prediction, receipt, and prediction error However, their study used an event-related design (which generally produces less robust activations than block designs), small monetary rewards, and psychiatrically healthy participants. In contrast, our ACR task, which uses a block design and alcohol cues among non-treatment-seeking alcohol dependent individuals, may evoke a more powerful and reliable striatal reward signal.

This study demonstrates the value of considering group vs. within-participants reliability, a consideration aided by use of HLM analysis techniques. While group-wise fMRI activation has demonstrated fair to good reliability across sessions for many task paradigms, within-participants activations are often highly variable (Aron et al., 2006). In the present study, there were no significant between-scan differences in the magnitude of group-wise activation in any of the anatomically defined ROIs. However, within-participants activations demonstrated good reliability only in the anatomically defined right VS and DS, and the functionally defined region encompassing right caudate and putamen (though reliability in right VS was driven somewhat by one participant who displayed low activation in both scans). Traditional repeated measures ANOVA can assess group-wise differences in activation (e.g., session, drug, or demographic variables), but HLM allows consideration of individual trajectories (i.e., changes in the magnitude of alcohol vs. beverage activation), as well as the influence of group-level variables on these trajectories. Critically, it models error variance associated with these different levels more accurately than ANOVA. In this regard, HLM represents a powerful method of analysis for serial block fMRI designs, which produce inherently hierarchically nested data.

These data also point to the importance of statistical comparison between serial fMRI images, particularly if these images are thresholded. Though activation was somewhat attenuated between scans in the anatomically defined ROIs, none of these differences was statistically significant at the group level. Similarly, while the left-sided functionally defined ROI from Scan 1 did not survive thresholding in Scan 2, there were no significant between-scan differences at the group level in the timecourse signal extracted from the Scan 1-

defined functional regions. Given the consistent slight attenuation between scans in the striatal cue-elicited signal observed here, treatment outcome studies that employ serial fMRI scans might mistake subtle signal shifts relative to a threshold (i.e., a signal falling just above vs. just below a specified threshold) as evidence of a treatment affecting ROI activation. Thus, such studies might be advised to use a summary statistic from an ROI (i.e., an HLM-derived parameter estimate) and to employ between-scan changes in activation among a placebo group as a statistical benchmark of the degree of "expected" attenuation above and beyond which a treatment group would have to demonstrate an effect.

It is unclear why activation was more stable in the right striatal ROIs than the left. In whole brain analyses, our ACR task has consistently elicited right-sided VS activation (Myrick et al., 2004, 2008, 2010), as has every other published ACR functional imaging study that has demonstrated striatal activation (though left-sided activations are also common). While structural and functional abnormalities in the left caudate have previously been reported among depressed individuals (e.g., Gabbay et al., 2007), there is a relative paucity of data on functional or structural striatal lateralization in alcohol dependence. Tiihonen and colleagues (1998) reported that a small subset of late-onset Type I alcoholic patients demonstrated markedly reduced presynaptic DA binding in the left caudate, but Sullivan et al. (2005) reported no asymmetrical volume reduction of caudate, putamen, or NAcc in a large sample of alcohol-dependent men, among whom all three structures were smaller relative to controls. It is possible that our data reflect normal right-left asymmetry of striatal DA receptors (e.g., Vernaleken et al., 2007); if cue-elicited activation is driven by DA signaling, and such signaling is normally lower in the left striatum, it may simply be more difficult to measure reliably. Alternatively, lower stability in the left striatum may be a function of the relatively small sample size and commensurately limited statistical power.

There were several limitations to this study. Though consistent with other recent test-retest fMRI studies (e.g., Aron et al., 2006; Freyer et al., 2009; Raemakers et al., 2007; Schunck et al., 2008), the sample size was relatively small, limiting the generalizability of results. Due to the small sample size, this study was also not powered to evaluate the stability of activation in other areas associated with alcohol cue reactivity, such as the anterior cingulate, ventromedial and dorsolateral prefrontal cortices, orbitofrontal cortex, and amygdala; future work should address this issue. The anatomical ROIs were defined on an anatomical template to which participants' individual functional scans were normalized, rather than on each participant's anatomical scan. While this approach is common for ROI analysis, it results in some lack of overlap between participants in the underlying tissue, because spatial normalization cannot perfectly match brains across individuals. However, our use of the Harvard-Oxford probabilistic atlas for ROI definition, rather than a single-subject atlas, follows best practices for this procedure (see Poldrack, 2007). Finally, the interval between scans was relatively short. We chose a two-week interval to avoid the possibility that participants would change their drinking patterns over a longer interval, thus potentially confounding any change in between-session activation. Further, this interval is pragmatic for pharmacological treatment trials because it allows most medication dosing to reach steady state and enables evaluation of participants before dropout rates increase (Mann et al., 2009; Myrick et al., 2008, 2010). However, future work should address the long-term stability of alcohol cue-elicited activation.

5. Conclusions

In summary, this small study demonstrated good within-participants reliability of striatal cue-elicited activation (especially in the right hemisphere) using HLM, an approach uniquely suited to the evaluation of fMRI serial block design data. However, further evaluation of the reliability of such activation in larger samples is necessary. If cue-elicited

striatal activation continues to demonstrate good within-participants reliability, it could ultimately represent a biomarker, or potential predictor, of treatment response for individual alcohol-dependent patients.

Acknowledgments

References

Anton RF, Moak DH, Latham PK. The obsessive compulsive drinking scale: A new method of assessing outcome in alcoholism treatment studies. Archives of General Psychiatry. 1996; 53:225–231. [PubMed: 8611059]

Aron AR, Gluck MA, Poldrack RA. Long-term test-retest reliability of functional MRI in a classification learning task. NeuroImage. 2006; 29:1000–1006. [PubMed: 16139527]

Atallah HE, Lopez-Paniagua D, Rudy JW, O'Reilly RC. Separate neural substrates for skill learning and performance in the ventral and dorsal striatum. Nature Neuroscience. 2007; 10:126–131.

Balleine BW, O'Doherty JP. Human and rodent homologies in action control: corticostriatal determinants of goal-directed and habitual action. Neuropsychopharmacology. 2010; 35:48–69. [PubMed: 19776734]

Beck A, Schlagenhauf F, Wüstenberg T, Hein J, Kienast T, Kahnt T, Schmack K, et al. Ventral striatal activation during reward anticipation correlates with impulsivity in alcoholics. Biological Psychiatry. 2009; 66:734–742. [PubMed: 19560123]

Berridge KC, Robinson TE. What is the role of dopamine in reward: hedonic impact, reward learning, or incentive salience? Brain Research. Brain Research Reviews. 1998; 28:309–369. [PubMed: 9858756]

Bohn MJ, Krahn DD, Staehler BA. Development and initial validation of a measure of drinking urges in abstinent alcoholics. Alcoholism, Clinical and Experimental Research. 1995; 19:600–606.

Braus DF, Wrase J, Grüsser S, Hermann D, Ruf M, Flor H, et al. Alcohol-associated stimuli activate the ventral striatum in abstinent alcoholics. Journal of Neural Transmission. 2001; 108:887–894. [PubMed: 11515754]

Caceres A, Hall DL, Zelaya FO, Williams SCR, Mehta MA. Measuring fMRI reliability with the intra-class correlation coefficient. NeuroImage. 2009; 45:758–768. [PubMed: 19166942]

Clément F, Belleville S. Test-retest reliability of fMRI verbal episodic memory paradigms in healthy older adults and in persons with mild cognitive impairment. Human Brain Mapping. 2009; 30:4033–4047. [PubMed: 19492301]

Filbey FM, Claus E, Audette AR, Niculescu M, Banich MT, Tanabe J, et al. Exposure to the taste of alcohol elicits activation of the mesocorticolimbic neurocircuitry. Neuropsychopharmacology. 2008; 33:1391–1401. [PubMed: 17653109]

Fliessbach K, Rohe T, Linder NS, Trautner P, Elger CE, Weber B. Retest reliability of reward-related BOLD signals. NeuroImage. 2010; 50:1168–1176. [PubMed: 20083206]

Freyer T, Valerius G, Kuelz A, Speck O, Glauche V, Hull M, Voderholzer U. Test-retest reliability of event-related functional MRI in a probabilistic reversal learning task. Psychiatry Research. 2009; 174:40–46. [PubMed: 19783412]

Gabbay V, Hess DA, Liu S, Babb JS, Klein RG, Gonen O. Lateralized caudate metabolic abnormalities in adolescent major depressive disorder: A proton MR spectroscopy study. American Journal of Psychiatry. 2007; 164:1881–1889. [PubMed: 18056244]

Genovese CR, Noll DC, Eddy WF. Estimating test-retest reliability in functional MR imaging. I: Statistical methodology. Magnetic Resonance in Medicine. 1997; 38:497–507. [PubMed: 9339452]

George DT, Gilman J, Hersh J, Thorsell A, Herion D, Geyer C, Peng X, et al. Neurokinin 1 receptor antagonism as a possible therapy for alcoholism. Science. 2008; 319:1536–1539. [PubMed: 18276852]

Grüsser SM, Wrase J, Klein S, Hermann D, Smolka MN, Ruf M, et al. Cue-induced activation of the striatum and medial prefrontal cortex is associated with subsequent relapse in abstinent alcoholics. Psychopharmacology. 2004; 175:296–302. [PubMed: 15127179]

Hariri AR, Brown SM, Williamson DE, Flory JD, de Wit H, Manuck SB. Preference for immediate over delayed rewards is associated with magnitude of ventral striatal activity. Journal of Neuroscience. 2006; 26:13213–13217. [PubMed: 17182771]

Havel P, Braun B, Rau S, Tonn J, Fesl G, Brückmann H, Ilmberger J. Reproducibility of activation in four motor paradigms. An fMRI study. Journal of Neurology. 2006; 253:471–476. [PubMed: 16283098]

Heinz A, Siessmeier T, Wrase J, Hermann D, Klein S, Grüsser SM, et al. Correlation between dopamine D(2) receptors in the ventral striatum and central processing of alcohol cues and craving. American Journal of Psychiatry. 2004; 161:1783–1789. [PubMed: 15465974]

Imperato A, Di Chiara G. Preferential stimulation of dopamine release in the nucleus accumbens of freely moving rats by ethanol. Journal of Pharmacology and Experimental Therapeutics. 1986; 239:219–228. [PubMed: 3761194]

Johnstone T, Somerville LH, Alexander AL, Oakes TR, Davidson RJ, Kalin NH, Whalen PJ. Stability of amygdala BOLD response to fearful faces over multiple scan sessions. NeuroImage. 2005; 25:1112–1123. [PubMed: 15850729]

Kareken DA, Claus ED, Sabri M, Dzemidzic M, Kosobud AEK, Radnovich AJ, et al. Alcohol-related olfactory cues activate the nucleus accumbens and ventral tegmental area in high-risk drinkers: preliminary findings. Alcoholism: Clinical and Experimental Research. 2004; 28:550–557.

Kelley AE. Ventral striatal control of appetitive motivation: role in ingestive behavior and reward-related learning. Neuroscience and Biobehavioral Reviews. 2004; 27:765–776. [PubMed: 15019426]

Maitra R, Roys SR, Gullapalli RP. Test-retest reliability estimation of functional MRI data. Magnetic Resonance in Medicine. 2002; 48:62–70. [PubMed: 12111932]

Mann K, Kiefer F, Smolka M, Gann H, Wellek S, Heinz A, et al. Searching for responders to acamprosate and naltrexone in alcoholism treatment: Rationale and design of the Predict Study. Alcoholism: Clinical and Experimental Research. 2009; 33:674–683.

Manuck SB, Brown SM, Forbes EE, Hariri AR. Temporal stability of individual differences in amygdala reactivity. American Journal of Psychiatry. 2007; 164:1613–1614. [PubMed: 17898358]

Melendez RI, Rodd-Henricks ZA, McBride WJ, Murphy JM. Alcohol stimulates the release of dopamine in the ventral pallidum but not in the globus pallidus: a dual-probe microdialysis study. Neuropsychopharmacology. 2003; 28:939–946. [PubMed: 12637946]

Monti PM, Binkoff JA, Abrams DB, Zwick WR, Nirenberg TD, Liepman MR. Reactivity of alcoholics and nonalcoholics to drinking cues. Journal of Abnormal Psychology. 1987; 96:122–126. [PubMed: 3584660]

Myrick H, Anton RF, Li X, Henderson S, Drobes D, Voronin K, George MS. Differential brain activity in alcoholics and social drinkers to alcohol cues: Relationship to craving. Neuropsychopharmacology. 2004; 29:393–402. [PubMed: 14679386]

Myrick H, Anton RF, Li X, Henderson S, Randall PK, Voronin K. Effect of naltrexone and ondansetron on alcohol cue-induced activation of the ventral striatum in alcohol-dependent people. Archives of General Psychiatry. 2008; 65:466–475. [PubMed: 18391135]

Myrick H, Li X, Randall PK, Henderson S, Voronin K, Anton RF. The effect of aripiprazole on cue-induced brain activation and drinking parameters in alcoholics. Journal of Clinical Psychopharmacology. 2010; 30:365–372. [PubMed: 20571434]

Noll DC, Genovese CR, Nystrom LE, Vazquez AL, Forman SD, Eddy WF, Cohen JD. Estimating test-retest reliability in functional MR imaging. II: Application to motor and cognitive activation studies. Magnetic Resonance in Medicine. 1997; 38:508–517. [PubMed: 9339453]

O'Doherty J, Dayan P, Schultz J, Deichmann R, Friston K, Dolan RJ. Dissociable roles of ventral and dorsal striatum in instrumental conditioning. Science. 2004; 304:452–454. [PubMed: 15087550]

Park SQ, Kahnt T, Beck A, Cohen MX, Dolan RJ, Wrase J, Heinz A. Prefrontal cortex fails to learn from reward prediction errors in alcohol dependence. Journal of Neuroscience. 2010; 30:7749–7753. [PubMed: 20519550]

Poldrack RA. Region of interest analysis for fMRI. Social Cognitive and Affective Neuroscience. 2007; 2:67–70. [PubMed: 18985121]

Potvin PJ, Schutz RW. Statistical power for the two-factor repeated measures ANOVA. Behavior Research Methods, Instruments, & Computers. 2000; 32:347–356.

Raemaekers M, Vink M, Zandbelt B, van Wezel RJA, Kahn RS, Ramsey NF. Test-retest reliability of fMRI activation during prosaccades and antisaccades. NeuroImage. 2007; 36:532–542. [PubMed: 17499525]

Raudenbush, SW.; Bryk, AS. Hierarchical Linear Models: Applications and Data Analysis Methods. 2nd ed. Newbury Park, CA: Sage; 2002.

Schultz W, Dayan P, Montague PR. A neural substrate of prediction and reward. Science. 1997; 275:1593–1599. [PubMed: 9054347]

Schunck T, Erb G, Mathis A, Jacob N, Gilles C, Namer IJ, Meier D, et al. Test-retest reliability of a functional MRI anticipatory anxiety paradigm in healthy volunteers. Journal of Magnetic Resonance Imaging. 2008; 27:459–468. [PubMed: 18307211]

Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. Psychological Bulletin. 1979; 86:420–428. [PubMed: 18839484]

Sinha R, Li CSR. Imaging stress- and cue-induced drug and alcohol craving: association with relapse and clinical implications. Drug and Alcohol Review. 2007; 26:25–31. [PubMed: 17364833]

Skinner HA, Allen BA. Alcohol dependence syndrome: measurement and validation. Journal of Abnormal Psychology. 1982; 91:199–209. [PubMed: 7096790]

Sobell, LC.; Sobell, MB. Timeline follow-back: A technique for assessing self-reported alcohol consumption. In: Allen, JP.; Litten, RZ., editors. Measuring alcohol consumption: Psychosocial and biochemical methods. Totowa, NJ: Humana Press; 1992. p. 41-72.

Specht K, Willmes K, Shah NJ, Jäncke L. Assessment of reliability in functional imaging studies. Journal of Magnetic Resonance Imaging. 2003; 17:463–471. [PubMed: 12655586]

Sullivan EV, Deshmukh A, De Rosa E, Rosenbloom MJ, Pfefferbaum A. Striatal and forebrain nuclei volumes: contribution to motor function and working memory deficits in alcoholism. Biological Psychiatry. 2005; 57:768–776. [PubMed: 15820234]

Tapert SF, Brown GG, Baratta MV, Brown SA. fMRI BOLD response to alcohol stimuli in alcohol dependent young women. Addictive Behaviors. 2004; 29:33–50. [PubMed: 14667419]

Tiffany ST. A cognitive model of drug urges and drug-use behavior: role of automatic and nonautomatic processes. Psychological Review. 1990; 97:147–168. [PubMed: 2186423]

Tiihonen J, Vilkman H, Räsänen P, Ryynänen OP, Hakko H, Bergman J, Hämäläinen T, et al. Striatal presynaptic dopamine function in type 1 alcoholics measured with positron emission tomography. Molecular Psychiatry. 1998; 3:156–161. [PubMed: 9577840]

Tzourio-Mazoyer N, Landeau B, Papathanassiou D, Crivello F, Etard O, Delcroix N, et al. Automated anatomical labelling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single subject brain. NeuroImage. 2002; 15:273–289. [PubMed: 11771995]

Vernaleken I, Weibrich C, Siessmeier T, Buchholz H, Rösch F, Heinz A, Cumming P, et al. Asymmetry in dopamine D(2/3) receptors of caudate nucleus is lost with age. NeuroImage. 2007; 34:870–878. [PubMed: 17174574]

Volkow ND, Wang G, Telang F, Fowler JS, Logan J, Childress A, Jayne M, et al. Cocaine cues and dopamine in dorsal striatum: mechanism of craving in cocaine addiction. Journal of Neuroscience. 2006; 26:6583–6588. [PubMed: 16775146]

Vollstädt-Klein S, Wichert S, Rabinstein J, Bühler M, Klein O, Ende G, Hermann D, et al. Initial, habitual and compulsive alcohol use is characterized by a shift of cue processing from ventral to dorsal striatum. Addiction. 2010; 105:1741–1749. [PubMed: 20670348]

Wrase J, Grüsser SM, Klein S, Diener C, Hermann D, Flor H, Mann K, et al. Development of alcohol-associated cues and cue-induced brain activation in alcoholics. European Psychiatry. 2002; 17:287–291. [PubMed: 12381499]

Wrase J, Schlagenhauf F, Kienast T, Wüstenberg T, Bermpohl F, Kahnt T, Beck A, et al. Dysfunction of reward processing correlates with alcohol craving in detoxified alcoholics. NeuroImage. 2007; 35:787–794. [PubMed: 17291784]

Yeh P, Gazdzinski S, Durazzo TC, Sjöstrand K, Meyerhoff DJ. Hierarchical linear modeling (HLM) of longitudinal brain structural and cognitive changes in alcohol-dependent individuals during sobriety. Drug and Alcohol Dependence. 2007; 91:195–204. [PubMed: 17644276]

Yoshimoto K, McBride WJ, Lumeng L, Li TK. Ethanol enhances the release of dopamine and serotonin in the nucleus accumbens of HAD and LAD lines of rats. Alcoholism, Clinical and Experimental Research. 1992; 16:781–785.

Zhang L, Doyon WM, Clark JJ, Phillips PEM, Dani JA. Controls of tonic and phasic dopamine transmission in the dorsal and ventral striatum. Molecular Pharmacology. 2009; 76:396–404. [PubMed: 19460877]
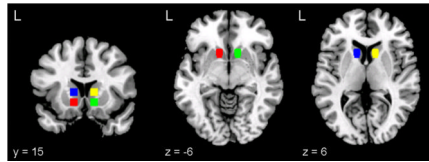
Figure 1.
Anatomically defined regions of interest displayed on the MNI single-subject T1 anatomical template: left (red) and right (green) ventral striatum and left (blue) and right (yellow) dorsal striatum.
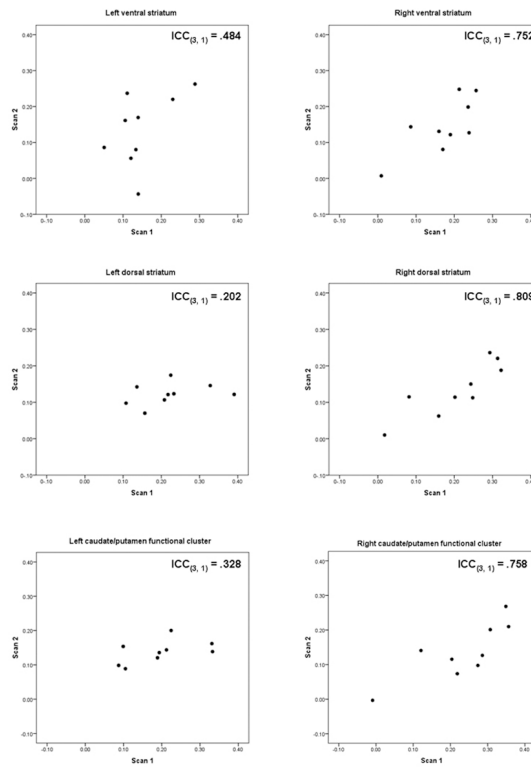
Figure 2.
Comparison between scans within the anatomically and functionally defined regions of interest. Points plotted are each participant's Scan 1 and Scan 2 HLM-derived parameter estimates (slopes) in arbitrary units for the alcohol vs. neutral beverage contrast. $ICC_{(3, 1)}$ = two-way mixed single measures intraclass correlation coefficient.
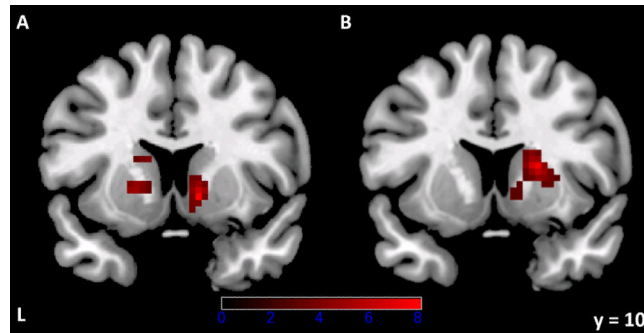
Figure 3.
Functionally defined clusters of alcohol vs. neutral beverage activation ($z > 1.96$, cluster-based $p < .05$, corrected for multiple comparisons, $k = 25$ voxels) within the striatum in Scan 1 (A) and Scan 2 (B). Images are neurologically oriented (i.e., the right side of the image corresponds to the right side of the brain) and are displayed on the same coronal slice. Colors correspond to t statistics for the alcohol vs. neutral beverage contrast.

Table 1

Demographic and drinking statistics for the final sample (*N* = 9).

	Mean (SD)
Age (years)	34.7 (12.4)
Gender (% male)	66%
Education (years)	15.6 (2.5)
Alcohol Dependence Scale	9.9 (5.5)
Obsessive Compulsive Drinking Scale	15.9 (5.0)
Baseline drinks/drinking day	5.7 (3.0)
Between-scans drinks/drinking day	4.8 (2.7)
Baseline % days abstinent	30.0 (18.2)
Between-scans % days abstinent	26.2 (17.2)
Scan 1 Alcohol Urge Questionnaire score	21.4 (4.4)
Scan 2 Alcohol Urge Questionnaire score	22.6 (5.2)

Table 2

Magnitude of alcohol vs. neutral beverage contrast in each scan, groupwise comparison between scans, and within-participants stability (ICC) of activation.

ROI	Scan 1	Scan 2	Scan 1 vs. Scan 2	$ICC_{(3, 1)}$
Anatomically defined regions of interest				
Right VS	$t(8) = 3.57, p = .009$	$t(8) = 2.58, p = .033$	$t(16) = -0.42, p = .68$	$.752, p = .006$
Right DS	$t(8) = 3.68, p = .007$	$t(8) = 2.54, p = .035$	$t(16) = -1.05, p = .31$	$.809, p = .002$
Left VS	$t(8) = 2.58, p = .033$	$t(8) = 2.33, p = .048$	$t(16) = -0.14, p = .89$	$.484, p = .078$
Left DS	$t(8) = 3.73, p = .007$	$t(8) = 2.46, p = .039$	$t(16) = -1.43, p = .17$	$.202, p = .288$
Functionally defined regions of interest				
Right caudate/putamen (71 voxels)	$t(8) = 3.99, p = .005$	$t(8) = 2.48, p = .038$	$t(16) = -1.34, p = .20$	$.758, p = .006$
Left caudate/putamen (136 voxels)	$t(8) = 3.76, p = .007$	$t(8) = 3.24, p = .013$	$t(16) = -0.95, p = .36$	$.328, p = .177$