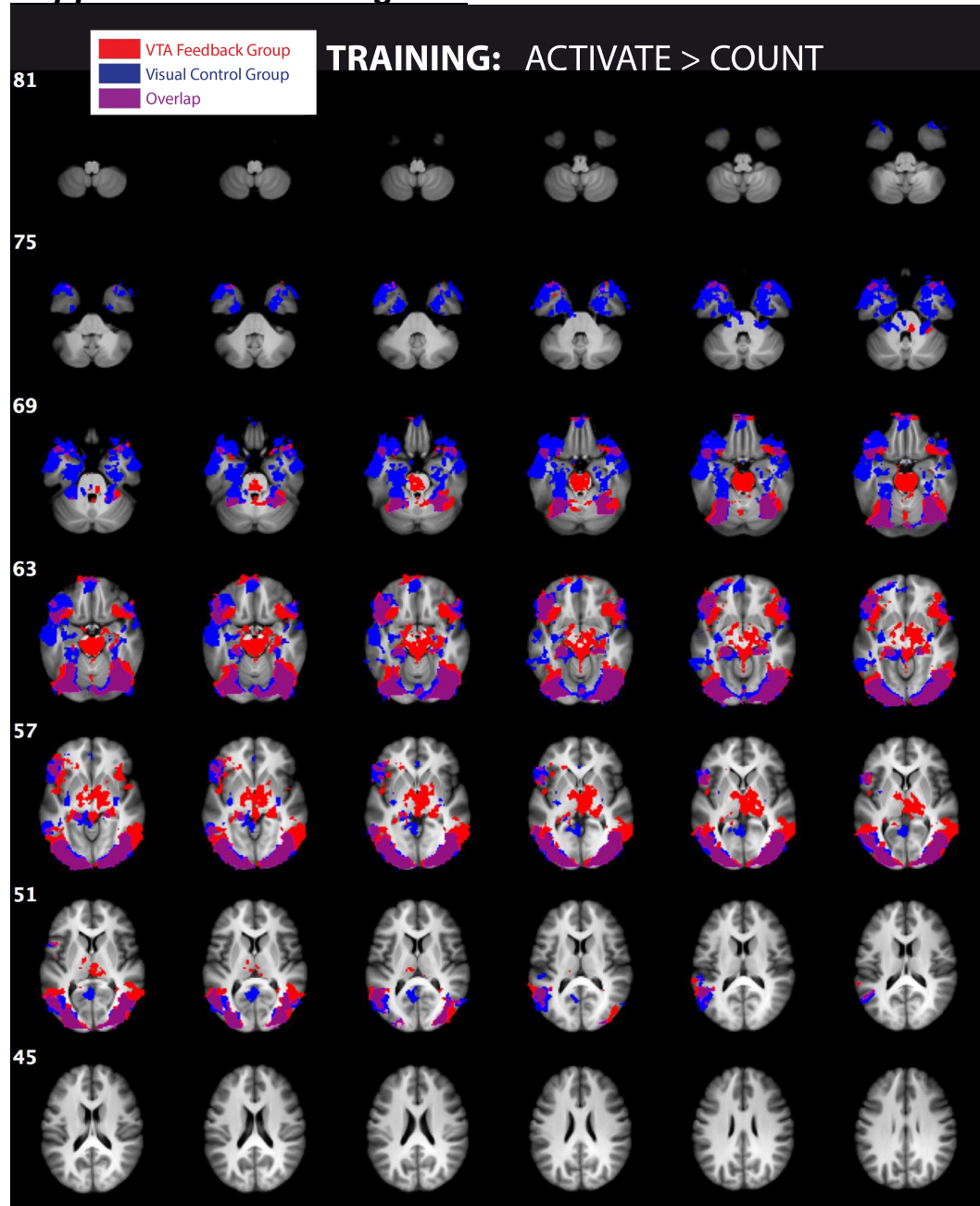


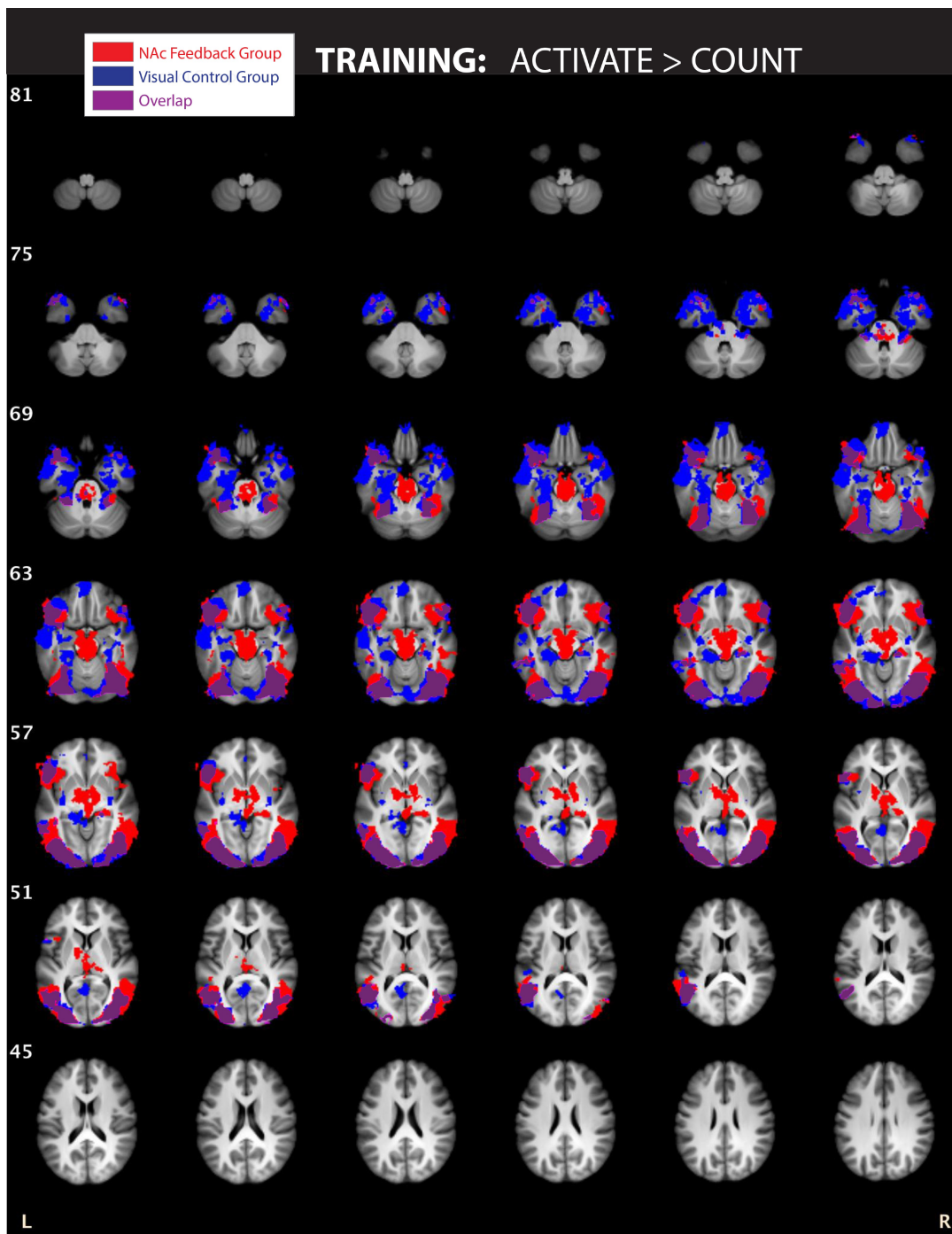
## **Inventory of Supplemental Information**

- **Supplemental Data**
  - Supplemental Figures
    - Figure S1: Whole-volume results for the comparison of ACTIVATE > COUNT during Training: VTA and Visual Control Feedback groups. Related to Figures 2 and 3
    - Figure S2: Whole-volume results for the comparison of ACTIVATE > COUNT during Training: NAcc and Visual Control Feedback groups. Related to Figures 4 and 5
    - Figure S3: Representative functional coverage from one participant. Related to Experimental Procedures: fMRI Acquisition
    - Figure S4: Comparison of event-related averages for veridical neurofeedback groups. Related to Figures 3 and 5
  - Supplemental Tables
    - Table S1: Comparison of self-reported Success, Difficulty, and Motivation ratings across groups during ACTIVATE Trials. Related to Experimental Procedures: Task and Instruction
    - Table S2: Analyses of Mesolimbic Network Functional Connectivity Changes from Pre-Test to Post-Test by Pathway. Related to Figure 6
    - Table S3: Analyses of Mesolimbic Network Functional Connectivity Changes from Pre-Test to Training by Pathway. Related to Figure 6 and Table S2
- **Supplemental Experimental Procedures**
  - Supplemental Methods
    - Participant Instruction
    - Within-Scan Subjective Ratings and Post-Scan Questionnaires
    - Within-Scan ROI localization and rt-fMRI Analysis
    - ROI Definitions
  - Supplemental Control Analyses
    - fMRI: a priori VTA ROI Control Analyses
      - Alternate Contrast – Activate > Rest
      - Thermometer Height Model (Training phase)
      - Positive Feedback Analysis (Training & Post-Test phases)
    - fMRI: a priori NAcc ROI Control Analyses
      - Temporal Dynamics of NAcc Responses During Neurofeedback
      - Replication of Greer, et al. (2014) NAcc Analyses
- **Supplemental References**

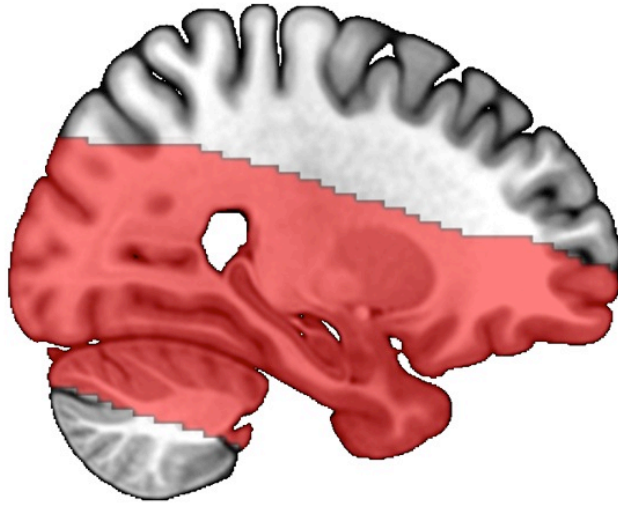
## Supplemental Data: Figures



**Figure S1. Whole volume results for the comparison of ACTIVATE > COUNT during Training: VTA and Visual Control Feedback groups. Related to Figures 2 and 3.** This figure illustrates regions engaged by the VTA Feedback group (red), Visual Control group (blue), and overlap between the groups (purple). Note the high overlap of regions engaged by both groups. See Figure S3 for slice coverage.

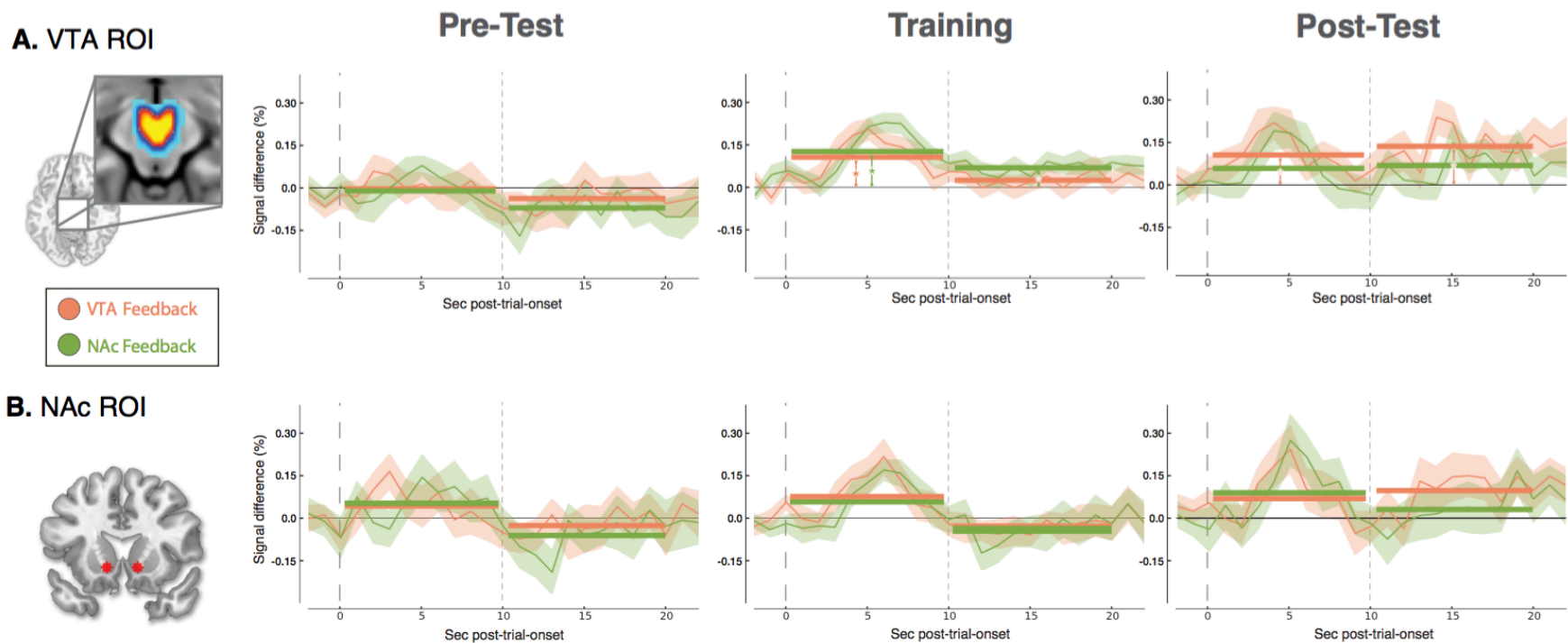


**Figure S2. Whole volume results for the comparison of ACTIVATE > COUNT during Training: NAcc and Visual Control Feedback groups. Related to Figures 4 and 5.** This figure illustrates regions engaged by the NAcc Feedback group (red), Visual Control group (blue), and overlap between the groups (purple). Note the high overlap of regions engaged by both groups. See Figure S3 for slice coverage.



**Figure S3. Representative functional coverage from one participant. Related to Experimental Procedures: fMRI Acquisition.** Using a 1 second TR, we acquired 18 slices aligned to include midbrain, striatal, medial temporal lobe, and ventral frontal regions.





**Figure S4. Comparison of event-related averages for NAcc and VTA veridical neurofeedback groups. Related to Figures 3 and 5.** Event-Related Average timecourses for ACTIVATE > COUNT during Test and Training trials. Waveforms represent % signal difference from baseline (**Orange – VTA Feedback group; Green – NAcc Feedback group**; shading, +/- s.e.m). The timecourse for both ACTIVATE and COUNT is calculated relative to the preceding 3-second inter-trial interval. To compare the timeseries, we subtracted COUNT from the ACTIVATE timeseries. Timecourses were segmented at 10 seconds to examine sustained activation (solid horizontal bars represent means). **A. Timecourses extracted from probabilistic VTA ROI, by feedback group.** In both the Pre-Test or the Post-Test, there were no group differences in the ability to activate the VTA, as revealed by no significant main effect of group ( $F(1, 37) = 1.08, P = 0.31$ ) on a 2 (run: Pre-Test, Post-Test) x 20 (timepoint: 1-20) x 2 (group: VTA Feedback, NAcc Feedback) ANOVA. There was a *significant main effect of run*, ( $F(1, 37) = 9.28, P < 0.005$ ), a *significant main effect of timepoint*

( $F(14.00, 517.83) = 3.66, P < 0.001$ ), a trend-level interaction of run x timepoint ( $F(15.58, 576.40) = 1.53, P = 0.09$ ) and no other significant interactions (all  $P_s > 0.47$ ). To examine Training phase activation (middle panel) we ran a 3 (run: Training Run 1, Training Run 2, Training Run 3) x 20 (timepoint: 1-20) x 2 (group: VTA Feedback, NAcc Feedback) ANOVA. There was a *significant main effect of timepoint* ( $F(13.95, 516.29) = 7.93, P < 0.001$ ), a trend-level interaction of run x group x timepoint ( $F(33.02, 1221.75) = 1.38, P = 0.07$ ), with no other significant main effects or interactions: group ( $F(1, 37) = 1.19, P = 0.28$ ), run ( $F(2, 74) = 0.48, P = 0.62$ ), interactions (all  $P_s > 0.50$ ). Post hoc t-tests to examine activation above baseline are as follows: Pre-Test: No significant activation above baseline ( $P_s \geq 0.1$ ). Training: The VTA Feedback group significantly activated the VTA above baseline during the early phase of the trial ( $P < 0.0005$ ). The NAcc Feedback group produced significant VTA activation during both early and late phases of the trial [early:  $P < 0.0001$ ; late:  $P < 0.01$ ]. Post-Test: The VTA Feedback group generated significant VTA activation relative to baseline over both phases of the trial [early:  $P < 0.05$ ; late:  $P < 0.005$ ]. The NAcc Feedback group did not produce VTA activation that significantly differed from baseline during the Post-Test ( $P_s \geq 0.1$ ).

**B. Timecourses extracted from the NAcc ROI, by feedback group.** There were no group differences in the ability to activate the NAcc in either the Pre-Test or the Post-Test as revealed by no significant main effect of group ( $F(1, 37) = 0.14, P = 0.71$ ) on a 2 (run: Pre-Test, Post-Test) x 20 (timepoint: 1-20) x 2 (group: VTA Feedback, NAcc Feedback) ANOVA. There was a *significant main effect of timepoint* ( $F(13.33, 493.26) = 3.88, P < 0.001$ ), a nearly significant main effect of run ( $F(1, 37) = 2.85, P = 0.10$ ), and no significant interactions (all  $P_s \geq 0.20$ ). To examine Training phase activations (middle panel) we ran a 3 (run: Training Run 1, Training Run 2, Training Run 3) x 20 (timepoint: 1-20) x 2 (group: VTA Feedback, NAcc Feedback) ANOVA. There was a *significant main effect of timepoint* ( $F(10.07, 372.60) = 7.13, P < 0.001$ ), but no other significant main effects or interactions: group ( $F(1, 37) = 0.08, P = 0.78$ ), run ( $F(2, 74) = 0.78, P = 0.46$ ), interactions (all  $P_s > 0.43$ ). Post hoc t-tests to examine activation above baseline were all non-significant, as follows: Pre-Test (all  $P_s > 0.20$ ), Training (non-significant after sequential Bonferroni correction;  $P_s \geq 0.03$ ), Post-Test (non-significant after sequential Bonferroni correction;  $P_s \geq 0.04$ ). There were no significant VTA Feedback vs. NAcc Feedback group differences in either VTA or NAcc activation during any stage of the task. Thus, although the groups did not significantly differ from each other, the NAcc feedback group showed no significant activations, while the group that received veridical VTA neurofeedback significantly activated the VTA above baseline (and relative to control groups) following training.

**Supplemental Data: Tables**

<u>Group 1</u>	<u>Group 2</u>	<u>Success</u>	<u>Difficulty</u>	<u>Motivation</u>
VTA	VC	<b><math>t(37) = 2.29, p &lt; 0.05</math></b> VTA $M = 4.5$ ; VC $M = 5.2$	<b><math>t(37) = 3.91, p &lt; 0.0005</math></b> VTA $M = 5.6$ ; VC $M = 4.1$	$t(37) = 0.25, p > 0.1$
VTA	FF	<b><math>t(31) = 2.48, p &lt; 0.05</math></b> VTA $M = 4.5$ ; FF $M = 3.4$	$t(31) = 0.42, p > 0.1$	<b><math>t(31) = 2.97, p &lt; 0.01</math></b> VTA $M = 6.1$ ; FF $M = 4.9$
VTA	NAcc	$t(37) = 0.55, p > 0.1$	$t(37) = 0.03, p > 0.1$	$t(37) = 0.53, p > 0.1$
NAcc	VC	<b><math>t(38) = 2.66, p &lt; 0.05</math></b> NAcc $M = 4.3$ ; VC $M = 5.2$	<b><math>t(38) = 3.80, p &lt; 0.001</math></b> NAcc $M = 5.6$ ; VC $M = 4.1$	$t(38) = 0.44, p > 0.1$
NAcc	FF	$t(32) = 1.90, p = 0.07$	$t(32) = 0.43, p > 0.1$	$t(32) = 1.81, p = 0.08$
<b>Bold denotes <math>p</math>-value <math>&lt; 0.05</math>; <math>M</math> = mean</b>				

**Table S1. Comparison of self-reported Success, Difficulty, and Motivation ratings across groups during ACTIVATE trials. Related to Experimental Procedures: Task and Instruction.** Following the scan session, all participants rated their subjective sense of motivation, excitement, success, and difficulty for the ACTIVATE trials. To assess differences in self-reported ratings across groups we performed a MANVOA using group as our independent variable (VTA Feedback, Visual Control [VC], NAcc Feedback, False Feedback [FF]) and the following dependent variables: success, difficulty, motivation, and excitement. A statistically significant MANOVA effect was obtained, Pillais' Trace = 0.47,  $F(12, 195) = 3.01, P < 0.01$ ; Box's M value was non-significant 45.51,  $P = 0.10$  indicating use of MANOVA was appropriate). Follow-up ANOVAs revealed significant group differences for Success ( $F(3, 66) = 6.29, P < 0.01$ ), Difficulty ( $F(3, 66) = 7.99, P < 0.001$ ), and Motivation ( $F(3, 66) = 3.41, P < 0.05$ ), but not Excitement ( $F(3, 66) = 0.88, P > 0.1$ ). Post hoc t-tests (reported in the table) revealed both the VTA and NAcc groups reported higher difficulty (Means 5.6) and lower success (Mean 4.5 and 4.3) ratings than the Visual Control group (Mean difficulty 4.1; success 5.2). The False Feedback group reported lower ratings of success (Mean 3.4) and motivation (4.9) than the VTA group (Mean success 4.5; motivation 6.1).

<u>Pathway</u>		<u>Main Effects</u>		<u>Interaction</u>	<u>Post Hoc T-Tests</u>	
<u>ROI 1</u>	<u>ROI 2</u>	<u>Test Run</u> (Post vs. Pre)	<u>Feedback Group</u> (NAcc vs. VTA)	<u>Test Run X Group</u>	<u>VTA Feedback group</u> <i>Post-Test &gt; Pre-Test</i>	<u>NAcc Feedback group</u> <i>Post-Test &gt; Pre-Test</i>
VTA	L-HPC	<b><math>F(1,37) = 5.18, p &lt; 0.05</math></b>	$F(1,37) = 2.89, p = 0.10$ †	$F(1,37) = 3.37, p = 0.07$ †	$t(18) = 2.68, p < 0.05$	$t(19) = 0.34, p > 0.1$
VTA	R-HPC	$F(1,37) = 0.92, p > 0.1$	<b><math>F(1,37) = 3.92, p = 0.06</math> †</b>	<b><math>F(1,37) = 4.13, p &lt; 0.05</math></b>	<b><math>t(18) = 2.18, p &lt; 0.05</math></b>	$t(19) = 0.74, p > 0.1$
VTA	L-CAUD	$F(1,37) = 1.06, p > 0.1$	$F(1,37) = 0.06, p > 0.1$	$F(1,37) = 2.46, p > 0.1$	—	—
VTA	R-CAUD	$F(1,37) = 0.10, p > 0.1$	$F(1,37) = 0.14, p > 0.1$	$F(1,37) = 2.13, p > 0.1$	—	—
VTA	NAcc	$F(1,37) = 0.22, p > 0.1$	<b><math>F(1,37) = 4.82, p &lt; 0.05</math></b>	$F(1,37) = 2.76, p > 0.1$	—	—
NAcc	L-HPC	<b><math>F(1,37) = 2.84, p = 0.10</math> †</b>	$F(1,37) = 1.70, p > 0.1$	<b><math>F(1,37) = 4.09, p = 0.05</math> †</b>	<b><math>t(18) = 2.74, p &lt; 0.05</math></b>	$t(19) = 0.23, p > 0.1$
NAcc	R-HPC	$F(1,37) = 0.53, p > 0.1$	$F(1,37) = 1.07, p > 0.1$	<b><math>F(1,37) = 5.47, p &lt; 0.05</math></b>	<b><math>t(18) = 2.24, p &lt; 0.05</math></b>	$t(19) = 1.11, p > 0.1$
NAcc	L-CAUD	<b><math>F(1,37) = 6.36, p &lt; 0.05</math></b>	$F(1,37) = 0.40, p > 0.1$	$F(1,37) = 1.0, p > 0.1$	—	—
NAcc	R-CAUD	$F(1,37) = 1.70, p > 0.1$	$F(1,37) = 0.16, p > 0.1$	$F(1,37) = 1.04, p > 0.1$	—	—

† denotes trend level  $p$ -value ( $.05 < p \leq 0.10$ ). Bold denotes significant and trend level comparisons.

**Table S2: Analyses of Mesolimbic Network Functional Connectivity Changes from Pre-Test to Post-Test by Pathway.**

**Related to Figure 6.** Repeated-measures 2 x 2 ANOVA (test run: Pre-Test, Post-Test) x (group: VTA Feedback, NAcc Feedback) and post hoc t-test results. Post hoc t-tests comparing Pre- to Post-Test connectivity changes were performed following a significant (or trend level) test run x group interaction.

<u>Pathway</u>		<u>Main Effects</u>		<u>Interaction</u>	<u>Post Hoc T-Tests</u>	
<u>ROI 1</u>	<u>ROI 2</u>	<u>Run</u> (Pre-Test vs. Training)	<u>Feedback Group</u> (NAcc vs. VTA)	<u>Run X Group</u>	<u>VTA Feedback group</u> <i>Training &gt; Pre-Test</i>	<u>NAcc Feedback group</u> <i>Training &gt; Pre-Test</i>
VTA	L-HPC	$F(1,37) = 10.82, p < 0.005$	$F(1,37) = 2.76, p = 0.10 †$	$F(1,37) = 3.86, p = 0.06 †$	$t(18) = 3.00, p < 0.01$	$t(19) = 0.87, p > 0.1$
VTA	R-HPC	$F(1,37) = 6.68, p < 0.05$	$F(1,37) = 4.02, p = 0.05 †$	$F(1,37) = 3.70, p = 0.06 †$	$t(18) = 2.69, p < 0.05$	$t(19) = 0.46, p > 0.1$
VTA	L-CAUD	$F(1,37) = 8.03, p < 0.01$	$F(1,37) = 0.10, p > 0.1$	$F(1,37) = 3.66, p = 0.06 †$	$t(18) = 2.29, p < 0.05$	$t(19) = 0.48, p > 0.1$
VTA	R-CAUD	$F(1,37) = 2.87, p = 0.10 †$	$F(1,37) = 0.02, p > 0.1$	$F(1,37) = 1.43, p > 0.1$	—	—
VTA	NAcc	$F(1,37) = 1.96, p > 0.1$	$F(1,37) = 3.61, p = 0.07 †$	$F(1,37) = 0.93, p > 0.1$	—	—
NAcc	L-HPC	$F(1,37) = 2.08, p > 0.1$	$F(1,37) = 0.62, p > 0.1$	$F(1,37) = 0.77, p > 0.1$	—	—
NAcc	R-HPC	$F(1,37) = 1.62, p > 0.1$	$F(1,37) = 0.19, p > 0.1$	$F(1,37) = 1.31, p > 0.1$	—	—
NAcc	L-CAUD	$F(1,37) = 9.42, p < 0.005$	$F(1,37) = 1.20, p > 0.1$	$F(1,37) = 3.51, p = 0.07 †$	$t(18) = 2.69, p < 0.05$	$t(19) = 0.54, p > 0.1$
NAcc	R-CAUD	$F(1,37) = 1.50, p > 0.1$	$F(1,37) = 0.51, p > 0.1$	$F(1,37) = 2.31, p > 0.1$	—	—

† denotes trend level  $p$ -value ( $.05 < p \leq 0.10$ ). Bold denotes significant and trend level comparisons.

**Table S3. Analyses of Mesolimbic Network Functional Connectivity Changes from Pre-Test to Training by Pathway. Related to Figure 6 and Table S2.** Repeated-measures 2 x 2 ANOVA (run: Pre-Test, Training) x (group: VTA Feedback, NAcc Feedback) and post hoc t-test results. Post hoc t-tests comparing Pre-Test to Training connectivity changes were performed following a significant (or trend level) run x group interaction.

## **Supplemental Experimental Procedures**

### **Methods - Participant Instruction**

#### **Feedback Group (VTA, NAcc, FF) Instruction**

*Welcome to the experiment. The purpose of today's experiment is to investigate whether individuals can train themselves to increase activity in certain brain regions believed to be involved in reward processing and motivation. To help address this question, we will be providing you with real-time feedback of on-going activity within these regions, allowing you to adjust your strategies to find a method that works best for you and produces the greatest activity.*

*The task you'll be doing inside the MRI will be to try to regulate your own brain activity. There will be two versions of this task, one called Test, and one called Train.*

*The Test task simply consists of two conditions, COUNT and ACTIVATE, each signified by different symbols. During the COUNT conditions you will see two circles, side by side, presented on the screen; during the ACTIVATE condition, you will see two squares, side by side. Each condition will last for 20 seconds. During the COUNT condition, you will be asked to count backwards from the number 300 in increments of 4. After this counting period, you'll see a plus sign on the screen for a variable length of time. Focus on the plus sign and wait for the next symbol to appear. During the ACTIVATE condition try your best to produce a mental state of heightened motivation. For example, try to encourage yourself that you can increase your own brain signal. It may help to think of this task as a fun game. Telling yourself positive phrases, such as "you can do it!" or "increase that signal!" or imagining personally relevant scenarios may be useful strategies for you. These are the sorts of mental states we'd like you to try to produce. Importantly, you don't have to limit yourself to these phrases. Feel free to try anything that you think will be motivating. What motivates you may be different than what motivates me, so you should think about things that will best motivate you personally.*

*Also, please try to remain as still as possible throughout the entire scan and try to keep your breathing rate as consistent as possible. Changes, even minor ones in either breathing rate or head/body position, may significantly degrade the quality of the data. Also, just so you know, the COUNT and ACTIVATE conditions will be presented in random order, so they will not always alternate.*

*This Test task will last about 4 minutes. After a short rest, will begin the Train session. Any questions about the Test task?*

*In the Train task you'll be doing essentially the same things, only now you'll receive real-time feedback indicating how effectively you are increasing your brain signal. There are three conditions in this task: COUNT, REST, and ACTIVATE. The instructions for COUNT and ACTIVATE are identical to what you did in the Test task, however the cues representing these conditions will be different during the training runs. As before, each condition will last for 20 seconds.*

*The COUNT condition will now be represented by ovals and the instructions are the same as in the Test task: when the ovals are on the screen you should count backwards from 300 in increments of 4.*

*During the REST conditions you will see the oval stimuli again, only now there will also be a thermometer present in the middle of the screen. The height of the thermometer will be jumping around every second or so. NOTE: THESE THERMOMETER VALUES ARE RANDOMLY DETERMINED AND DO NOT REFLECT ON-GOING ACTIVITY IN YOUR BRAIN. When the thermometer level is above zero it will be yellow, and when it's below zero it will be green. During this condition you should simply rest and try not to think of anything in particular.*

*During the ACTIVATE conditions you will see two upward-facing arrows, as well as a thermometer in the center of the screen. However, on these trials, the thermometer is reporting on-going activity within brain*



areas involved in reward and motivation. Your task is to try to get – and keep – the thermometer as high as possible. Try to do the same things you did during the Test task. That is, try to encourage yourself to increase your brain signal as much as possible. When the thermometer level is above zero it will be red, and when it's below zero it will be blue. Since you are receiving feedback indicating how well you are doing, you can directly try to move the height of the thermometer with your brain. Try to encourage yourself to move that bar! You can adapt your strategies to discover which methods produce the greatest amount of activity. We ask that you stick with one strategy for the entire trial and if you find that strategy to be ineffective, feel free to try something else on the next trial.

While attempting to increase the thermometer height, there are a couple of things to keep in mind. First, the brain activity we are trying to measure is inherently noisy, meaning you might see the thermometer jump back and forth a bit. Don't get discouraged; try to maintain your focus and keep the average height of the thermometer as high as possible. Secondly, there may be a lag time of as much as 5 seconds between your current mental state and the feedback you see reflected on the thermometer. Be persistent with your strategies and don't give up on them too soon.

You may also notice the range of the thermometer adjusting over time. This is normal. Initially the thermometer will attempt to calibrate itself to the appropriate range for your brain activity.

To help motivate you, during the *ACTIVATE* conditions you will see 2 horizontal lines displayed on the thermometer (see task diagram instructions). The solid line indicates the threshold we'd like you to try and achieve. The dashed line indicates your average performance during the previous *ACTIVATE* condition. See if you can outperform your previous attempt! Additionally, you will see a value presented on the screen at the end of each trial indicating how well you performed on average for that trial.

After the experiment has concluded, we will ask you what strategies you thought worked best for you, so try to remember what seems to be working and what doesn't. Each Train session will take approximately 6 minutes. You will do three sessions.

The last thing we'll do in the experiment is a second Test task without feedback. The purpose of this session is to see how well you can activate your brain's reward and motivation regions now that you have had some training. You should think about which strategies worked best for you in the training session and use those strategies in this Test session. This session will be exactly like the first Test. Just as a reminder, there will be two conditions, a *COUNT* condition (circles) and an *ACTIVATE* condition (squares). You will not see the thermometer at any time point in this session.

At the end of every run you will see a rating screen asking how you did on each of the conditions. Do you have any questions about anything we've discussed?

## **Visual Control Group Instruction**

Welcome to the experiment. The purpose of today's experiment is to investigate whether individuals can train themselves to increase activity in certain brain regions believed to be involved in reward processing and motivation. To help address this question, we will be conducting some sessions in which you will try to increase activity within this brain region. As you perform the task, you may adjust your strategies to find a method that works best for you.

The task you'll be doing inside the MRI will be to try to regulate your own brain activity. There will be two versions of this task, one called Test, and one called Train.

The Test task simply consists of two conditions, *COUNT* and *ACTIVATE*, each signified by different symbols. During the *COUNT* conditions you will see two circles, side by side, presented on the screen; during the *ACTIVATE* condition, you will see two squares, side by side. Each condition will last for 20 seconds. During the *COUNT* condition, you will be asked to count backwards from the number 300 in

increments of 4. After this counting period, you'll see a plus sign on the screen for a variable length of time. Focus on the plus sign and wait for the next symbol to appear. During the ACTIVATE condition try your best to produce a mental state of heightened motivation. For example, try to encourage yourself that you can increase your own brain signal. It may help to think of this task as a fun game. Telling yourself positive phrases, such as "you can do it!" or, "increase that signal!" or imagining personally relevant scenarios may be useful strategies for you. These are the sorts of mental states we'd like you to try to produce. Importantly, you don't have to limit yourself to these phrases. Feel free to try anything that you think will be motivating. What motivates you may be different than what motivates me, so you should think about things that will best motivate you personally.

Also, please try to remain as still as possible throughout the entire scan and try to keep your breathing rate as consistent as possible. Changes, even minor ones in either breathing rate or head/body position, may significantly degrade the quality of the data. Also, just so you know, the COUNT and ACTIVATE conditions will be presented in random order, so they will not always alternate.

This Test task will last about 4 minutes. After a short rest, will begin the Train session. Any questions about the Test task?

In the Train task you'll be doing essentially the same things. There are three conditions in this task: COUNT, REST, and ACTIVATE. The instructions for COUNT and ACTIVATE are identical to what you did in the Test task, however the cues representing these conditions will be different during the training runs. As before, each condition will last for 20 seconds.

The COUNT condition will now be represented by ovals and the instructions are the same as in the Test task: when the ovals are on the screen you should count backwards from 300 in increments of 4.

During the REST condition you will see the oval stimuli again, only now there will also be a thermometer present in the middle of the screen. The height of the thermometer will display a repeating increasing and decreasing pattern. NOTE: THESE THERMOMETER VALUES ARE NOT MEANINGFUL. When the thermometer level is above zero it will be yellow, and when it's below zero it will be green. During this condition you should simply rest and try not to think of anything in particular.

During the ACTIVATE condition you will see two upward-facing arrows, as well as a thermometer in the center of the screen. Like in the REST condition, during these trials, the thermometer will display a repeating increasing and decreasing pattern. When the thermometer level is above zero it will be red, and when it's below zero it will be blue. Your task is to increase your brain signal as much as possible. We are interested in understanding how individuals can increase their brain activity in the presence of this mildly distracting stimulus. Try to do the same things you did during the Test task. You can adapt your strategies to discover which method you think causes the greatest amount of activity. We ask that you stick with one strategy for the entire trial and if you find that strategy to be ineffective, feel free to try something else on the next trial.

After the experiment has concluded, we will ask you what strategies you thought worked best for you, so try to remember what seems to be working and what doesn't.

Each Train session will take approximately 6 minutes. You will do three sessions.

The last thing we'll do in the experiment is a second Test task. The purpose of this session is to see how well you can activate your brain's reward and motivation regions now that you have had some training. You should think about which strategies worked best for you in the training session and use those strategies in this session. This session will be exactly like the first session. Just as a reminder, there will be two conditions, a COUNT condition (circles) and an ACTIVATE condition (squares). You will not see the thermometer at any time point in this session.

*At the end of every run you will see a rating screen asking how successful you felt you were in each of the conditions.*

*Do you have any questions about anything we've discussed?*

## **Methods - Within-Scan Subjective Ratings and Post-Scan**

### **Questionnaires:**

**Subjective ratings.** Following each run, participants were asked to rate their success at generating a motivated state in the ACTIVATE trials, counting backwards in the COUNT trials, and resting in the REST trials.

**Questionnaires.** We designed a questionnaire to probe participants' strategies and general experience during the task. Participants described strategies used to self-activate the target ROI (VTA or NAcc) and identified the most effective one(s). They rated their subjective success, motivation, excitement, and difficulty on a seven point Likert scale for ACTIVATE trials (see Table S1). In addition, participants completed the following standardized questionnaires after the scanning session: Behavioral-Inhibition Behavioral-Activation Scale, Temporal Experience of Pleasure Scale, State-Trait Anxiety Inventory, and Beck Depression Inventory.

### **Methods - Within-Scan ROI localization and rt-fMRI Analysis:**

Immediately following reconstruction, each EPI slice image was transmitted to a dedicated real-time analysis machine via a TCP/IP socket connection (Voyvodic, 1999). The real-time analysis machine ran custom python-based software that received incoming slice data, incrementally assembled a 4D data matrix, and calculated ROI-based statistical summaries on each completed 3D time point. The feedback source ROIs (See below and Ballard et al., 2011; Murty et al., 2014, and Greer et al., 2014) were transformed from MNI space to participant-specific functional space using the participant's high-resolution anatomical image and initial resting state functional run as references. During the real-time Training runs, the VTA (and NAcc) ROIs were used to compute a weighted (and non-weighted) average of raw voxel intensities as each new time point was acquired. The output from this calculation was made available via an additional TCP/IP socket connection to a separate machine controlling the experimental presentation. In this manner the thermometer display on the presentation machine was updated ~1/second.

### **Methods - ROI Definitions:**

**VTA ROI:** The VTA was defined on the basis of a custom probabilistic atlas (Ballard et al., 2011; Murty et al., 2014). Atlas available at [www.adcocklab.org](http://www.adcocklab.org).

**NAcc ROI:** An NAcc mask defined as 10 mm spheres centered at +/- 10, 14, -2 (following Greer et al., 2014; MNI coordinates taken from Knutson & Greer, 2008). The MNI space NAcc mask was transformed to functional space for each participant, thresholded at values >50% and binarized.

**Control ROI:** We selected a gray matter control region in Heschl's gyrus, which exhibits sparse VTA/SN innervation (Lewis et al., 2001) and minimal *a priori* task involvement. We defined this region using the Harvard-Oxford Cortical Structural Atlas, thresholded at probabilities > 50%, and binarized to create a mask.

**HPC ROI:** Separate ROIs were defined for the left and right hippocampus. For each ROI, the hippocampus was defined using an anatomical mask in MNI space cataloged within WFU pickatlas (Maldjian et al., 2003). The MNI space mask was transformed to functional space for each participant, thresholded at values >50% and then binarized.

**Caudate ROI:** Separate ROIs were defined for the left and right caudate nucleus. For each ROI, the caudate was defined using an anatomical mask in MNI space cataloged within WFU pickatlas (Maldjian et al., 2003). The MNI space mask was transformed to functional space for each participant, thresholded at values >50% and then binarized.

## **Control Analyses – fMRI: *a priori* VTA ROI analyses**

### ***Alternate contrast with REST replicates ACTIVATE Training effects in VTA***

To ensure that increased VTA activation in ACTIVATE relative to COUNT trials (**shown in Figures 2 and 3**) was not due to differences across conditions in mental imagery, internal speech, or memory specific to the COUNT condition, we conducted additional analyses. Specifically, we tested whether the ACTIVATE > REST comparison would replicate our ACTIVATE > COUNT results. Individuals often engage in mind-wandering during rest conditions with no specific instructed or implicit task demands (Mason et al., 2007). Therefore ACTIVATE and REST trials could be more similar in terms of engaged cognitive resources than ACTIVE and COUNT trials. To test this we performed a 3 (run: Training 1, 2, 3) x 2 (group: VTA Feedback, Visual Control) x 20 (timepoint: 1-20) ANOVA for the comparison of ACTIVATE versus REST.

Results from this model replicated the COUNT analyses: VTA activation was significantly greater in the VTA Feedback relative to the Visual Control group, evidenced by a main effect of group ( $F(1,37) = 8.56, P < 0.01$ ; follow-up post hoc t-test ( $t(37) = 2.93, P < 0.01$ )). There was also a main effect of timepoint ( $F(13.11, 484.92) = 9.06, P < 0.001$ ), but no effect of run ( $F(2, 74) = 0.43, P > 0.1$ ) and no significant interactions (all p values > 0.1). The VTA Feedback group displayed greater VTA activation than the Visual Control group in both the first and second halves of the trial [early: ( $t(37) = 3.21, P < 0.005$ ); late: ( $t(37) = 2.37, P < 0.05$ ), collapsing across runs].

### ***Thermometer height model – Training Phase***

The enhanced VTA activation observed in the VTA Feedback group during training (**shown in Figure 3**) could reflect a “reward” response to the height of the thermometer. According to this interpretation, the current value of the thermometer (which was normalized to occupy the same graphical range across subjects) should have influenced subsequent VTA activation, in the VTA Feedback group only. To test this alternative, we modeled the height of the thermometer over time for each individual. For each timepoint, the current thermometer value (positive or negative) was convolved with a canonical HRF and included as a regressor of interest in a GLM. The critical contrasts were between BOLD responses to ACTIVATE trials in which the thermometer showed veridical feedback (VTA Feedback group) or predictable values (Visual Control group) versus REST trials [in which the thermometer displays mimicked the ACTIVATE condition, (random for the VTA Feedback group and predictable for Visual Control group)].

If the thermometer display drove subsequent VTA activation, this would be demonstrated by differences in trial type (e.g., significant effects for ACTIVATE, but not REST trials), group differences (e.g., significant effects for VTA Feedback, but not Visual Control participants) or an interaction of group x trial type (e.g., significant effects for ACTIVATE trials in the VTA Feedback group only). To test this we performed a 2 (trial type: REST, ACTIVATE) x 2 (group: VTA Feedback, Visual Control) ANOVA. No significant results were observed: there was no main effect of trial type ( $F(1, 37) = 0.64, P = 0.43$ ), no main effect of group ( $F(1, 37) = 2.50, P = 0.12$ ), and no significant interaction of trial type x group ( $F(1, 37) = 0.02, P = 0.88$ ). These null results provide evidence that the thermometer display itself did not significantly predict subsequent VTA

activation, in either the VTA Feedback or control groups. Our results are thus inconsistent with brief phasic responses to any external task events, and suggest instead that activation was enhanced throughout the trial by internal representations.

### ***Positive Feedback analyses – Training phase***

We examined whether positive feedback drove the significant VTA activation we observed during the Post-Test (**shown in Figures 2 and 3**). The average percent of time (e.g., 64%) the thermometer was positive (above 0) during training predicted neither Post-Test activation ( $r(17) = 0.15, P = 0.55$ ) nor the improvement from Pre- to Post-Test activation (Post-Test minus Pre-Test VTA activation;  $r(17) = 0.09, P = 0.70$ ) in the VTA Feedback group. Nor was there a significant relationship between the number of trials on which the average feedback was positive and Post-Test activation ( $r(17) = 0.08, P = 0.74$ ) or improvement from Pre to Post-Test ( $r(17) = 0.09, P = 0.72$ ). Moreover, the VTA Feedback group reported significantly less subjective success than the Visual Control group (**see Table S1**), making the alternative interpretation that positive feedback drove our effects even less compelling.

### **Control Analyses – fMRI: a priori NAcc ROI analyses**

#### ***Temporal dynamics of NAcc responses during neurofeedback***

We performed fully corrected, post hoc comparisons examining the early and late phases of the trial to assess sustained activation (**shown in Figure 5**). BOLD signal was not significantly different from baseline for any group [NAcc Feedback: *early*  $t(19) = 1.79, P = 0.09$ , *late*  $t(19) = 0.82, P > 0.10$ ; Visual Control: *early*  $t(19) = 0.26, P > 0.10$ , *late*  $t(19) = 0.87, P > 0.10$ ; False Feedback: *early*  $t(13) = 1.20, P > 0.10$ , *late*  $t(13) = 1.62, P > 0.10$ ]. Thus, unlike in the VTA, no group succeeded in activating the NAcc during training.

#### ***Replication of Greer et al. (2014) NAcc Analyses***

Our analyses focused on different aspects of the NAcc response than what Greer et al. (2014) reported. Because we were focused on a sustained duration of response, we averaged over multiple time points - either the entire 20-second timecourse of each trial, or the 10-second early and late halves -- and computed statistics on these mean values (**shown in Figure 5**). Greer et al. (2014) focused on the magnitude of the response, and thus selected the peak TR of activation within the timecourse and computed statistics on those values. Since the mean is necessarily less than the peak, it is no surprise that our reported activations are less significant than what Greer et al. (2014) report. In the interest of replicating their findings, we recomputed our results using peak values instead of means. Using this measure, like Greer and colleagues, we found significant NAcc activation above baseline throughout all phases of the experiment [Pre-Test:  $t(19) = 3.65, p < 0.005$ ; Training:  $t(19) = 8.26, p < 0.0001$ ; Post-Test:  $t(19) = 3.54, p < 0.005$ ]. Unlike Greer et al. (2014), we failed to observe a significant increase in NAcc peak magnitude from the Pre-Test to Training phase. This discrepancy could be due to differences in participant instructions and strategy use (motivational strategies vs. positive affect/arousal).

### **Supplemental References**

Knutson, B., & Greer, S.M. (2008). Anticipatory affect: neural correlates and consequences for choice. *Philosophical Transactions of the Royal Society of London Series B, Biological Sciences*, 363(1511), 3771–3786.

Lewis, D.A., Melchitzky, D.S., Sesack, S.R., Whitehead, R.E., Sungyoung Auh, and Sampson, A. (2001). Dopamine transporter immunoreactivity in monkey cerebral cortex: Regional, laminar, and ultrastructural localization. *J. Comp. Neurol.* 432, 119–136.

Maldjian, J.A., Laurienti, P.J., Kraft, R.A., and Burdette, J.H. (2003). An automated method for neuroanatomic and cytoarchitectonic atlas-based interrogation of fMRI data sets. *NeuroImage* 19, 1233–1239.

Mason, M.F., Norton, M.I., Van Horn, J.D., Wegner, D.M., Grafton, S.T., and Macrae, C.N. (2007). Wandering minds: the default network and stimulus-independent thought. *Science* 315, 393–395.