Supplemental information (SI) for:
Neural mechanisms of attitude change toward stigmatized individuals: Temporoparietal junction activity predicts bias reduction. Yoona Kang and Emily B. Falk, Annenberg School for Communication, University of Pennsylvania

SI1. Demographic characteristics of participants with eligible data by condition

Participants in the lovingkindness and control conditions did not significantly differ with respect to age, gender, ethnicity, or education (*p*s>.10), with the exception of age in analyses linking neural to behavioral data (Table SI1). When we controlled for age in all analyses linking neural to behavioral data, the main results remained parallel.

Table SI1

	Analyses of behavioral data			**Analyses linking neural to behavioral data**		
	Loving kindness (*n*= 36)	Control (*n*=74)	Statistic (*p*)	Loving kindness (*n*=27)	Control (*n*= 68)	Statistic (*p*)
Age (yrs)	32.31 (11.83)	35.84 (12.17)	*F* = 2.08 (.15)	28.52 (9.85)	35.84 (12.07)	*F* = 7.84 (.01)
Female	21 (58.3%)	51 (68.9%)	χ^2= 0.78 (0.38)	18 (66.7%)	47 (69.1%)	χ^2= 0.00 (1.00)
Black	17 (47.2%)	30 (40.5%)	χ^2= 0.21 (.65)	11 (40.7%)	29 (42.7%)	χ^2< 0.01 (1.00)
Education	15.47 (2.55)	16.16 (3.13)	*F* = 1.30 (.26)	15.85 (3.00)	16.04 (3.27)	*F* = 0.07 (.80)
Implicit bias (*d*)	0.59 (0.40)	0.60 (0.41)	*F* = 0.02 (.90)	0.63 (0.36)	0.64 (0.39)	*F* = 0.03 (.87)

Note: Mean values and sample sizes are displayed with standard deviations and percentages, respectively, in parentheses where applicable.

SI2. ROI analyses comparing activity during the intervention tasks

Activity in the RTPJ and VS during the lovingkindness and control tasks was

compared to test whether those in the lovingkindness condition show greater activity in the mentalizing and reward-related regions, compared to those in the control condition. The lovingkindness task compared to control task elicited significantly greater activity in RTPJ, R^2=.059, β=.243, t(101)=2.516, p=.013, 95% CI [0.004, 0.036]. However, the VS activity did not differ across conditions, R^2=.005, β=.074, t(101)=0.741, p=.460, 95% CI [-0.009, 0.021].

SI3. Whole-brain analysis identifying regions associated with intervention tasks

Paralleling the findings reported by Kang et al., 2018, we re-ran these analyses with only participants who had complete data eligible for the current investigation. Consistent with the original report on these data, three separate contrasts were created for the whole-brain analyses: 1) a within-subjects contrast in the lovingkindness task when people made wishes for others vs. wishes for everyday activities (n=31), 2) a within-subjects contrast in the control task when people reflected on their lowest value situations vs. everyday activities (n=72), and 3) a between-subjects contrast of lovingkindness vs. control conditions (n=103). Whole-brain analyses were corrected for multiple comparisons using a family-wise false discovery rate (FDR) with corrected p value of .05 and cluster-corrected at k=10. For the between-subjects results contrasting lovingkindness vs. control conditions, no cluster survived the FDR correction, and Figure SI3c and Figure SI3d show results at p<.005, k=10. Both the lovingkindness (Figure SI3a) and control (Figure SI3b) trials elicited greater activity within the regions associated with mentalizing (e.g., RTPJ, VMPFC, MMPFC, PC) and reward processing (e.g., VS, VMPFC), compared to everyday activity control trials. However, when the lovingkindness and control conditions were directly contrasted, activity in these regions were generally stronger for the lovingkindness condition (Figure SI3c). See Table SI3 below for complementary whole-brain analysis results.

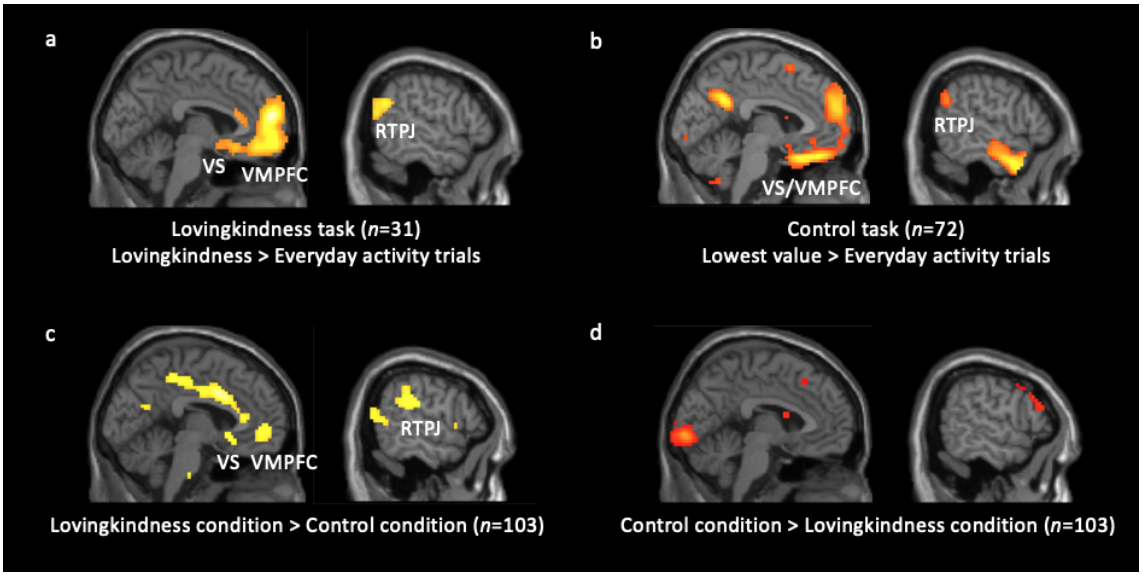

Figure SI3 a. A within-subjects contrast of lovingkindness task comparing lovingkindness > everyday activity trials; **b**. A within-subjects contrast of control task comparing lowest value > everyday activity trials; **c**. A between-subjects contrast of lovingkindness condition > control condition; **d**. A between-subjects contrast of control condition > lovingkindness conditions. Only positive clusters are displayed.

Table SI3. Whole-brain results showing regions with increased activity during the intervention tasks, $p<.05$ FDR corrected, $k=10$ for within-subjects contrasts (lovingkindness, control tasks) and $p<.005$, $k=10$, for the between-subjects contrast (lovingkindness > control condition). Note: L = left; R = right; Peak voxels and local maxima are reported in clusters that extend across the ventromedial prefrontal cortex and ventral striatum.

Region	*x*	*y*	*z*	*size*	*t*
Lovingkindness (*n*=31)					
Well-wishes > Everyday activity trials					
ventromedial prefrontal cortex	-9	56	16	1415	5.90
/ventral striatum	-6	20	-14	-	3.08
L temporal pole	-66	-13	-11	81	4.42
R temporal parietal junction	51	-58	25	182	4.40
Precuneus	0	-58	28	151	4.25
L temporal parietal junction	-45	-70	22	164	4.13
L superior frontal gyrus	-21	38	37	36	3.91
R inferior frontal gyrus	45	11	28	31	-3.74
L middle temporal gyrus	-51	-58	-14	52	-4.00
L superior parietal lobule	33	-64	43	144	-4.07
L middle frontal gyrus	-24	-4	49	148	-4.25
R middle frontal gyrus	24	32	-20	43	-4.53
R ventrolateral prefrontal cortex	48	53	-2	141	-4.98
R dorsolateral prefrontal cortex	51	26	40	355	-5.34
L dorso/ventrolateral prefrontal cortex	-42	5	31	784	-5.36
L superior parietal lobule	-30	-64	49	637	-5.55
occipital poles	-15	-100	-5	1618	-10.59
Control (*n*=72)					
Lowest value > Everyday activity trials					
L temporal pole	-48	14	-32	1641	8.27
R temporal pole	45	20	-32	885	7.79
Orbitofrontal cortex	0	29	-26	2045	6.65
/ventromedial prefrontal cortex	-12	56	28	-	6.29

precuneus	-6	-52	31	317	5.91
R occipital pole	39	-91	4	780	4.99
L occipital pole	-36	-94	1	1213	4.85
L postcentral gyrus	-54	-16	58	24	4.79
R temporoparietal junction	51	-58	31	105	4.63
R cerebellum	30	-79	-38	86	3.91
supplementary motor area	-6	11	61	52	3.86
R middle frontal gyrus	36	20	28	46	3.49
R cuneus	18	-91	43	15	3.44
L cerebellum	-6	-58	-47	16	3.18
caudate	6	11	13	12	3,08
L dorsomedial prefrontal cortex	-12	35	52	19	2.99
L precentral gyrus	-27	-19	61	61	-3.25
R insula	42	-10	-5	39	-3.39
L superior temporal gyrus	-21	8	-38	11	-3.44
L inferior frontal gyrus/insula	-63	8	22	214	3.76
L inferior frontal gyrus	-15	8	-20	31	-3.76
R precuneus	39	-76	37	66	-4.14
R Inferior parietal lobule	60	-37	40	162	-4.20
R inferior temporal gyrus	66	-52	-11	250	-4.38
L superior frontal gyrus	-21	11	52	232	-4.68
R cerebellum	27	-28	-35	491	-4.74
mid cingulate cortex	6	-4	28	244	-4.90
R posterior cerebellum	36	-61	-44	89	-5.01
Inferior frontal gyrus	45	41	7	203	-5.09
L middle temporal gyrus	-54	-58	-5	789	6.34
L inferior frontal gyrus	-42	41	10	394	-6.37
L middle frontal gyrus	-27	38	-14	214	-7.79
posterior cingulate cortex	-6	-34	43	3586	-8.24

Lovingkindness > Control (*n*=102)

R lingual gyrus	12	-70	-5	118	4.37
ventromedial prefrontal cortex/	-9	53	1	336	4.24
/ventral striatum	-6	20	-14	-	3.91
L Inferior temporal gyrus	-60	-34	-35	89	4.23
L middle temporal gyrus	-48	-79	22	193	3.94
L middle frontal gyrus	-21	26	43	126	3.93
cingulate gyrus	0	8	34	556	3.87
L precuneus/posterior cingulate gyrus	-15	-58	16	132	3.71
R superior frontal gyrus	27	35	37	92	3.53
L precuneus	-12	-52	76	31	3.52

R temporoparietal junction	63	-37	40	166	3.51
R cerebellum	33	-37	-47	44	3.45
L temporoparietal junction	-63	-40	34	92	3.43
R middle temporal gyrus	54	-67	16	130	3.37
L middle temporal gyrus	-60	-70	4	40	3.32
R insula	48	8	1	58	3.15
L middle frontal gyrus	-24	-4	49	14	-3.10
L middle temporal gyrus	-51	-40	-2	44	-3.29
L superior temporal gyrus	-48	14	-17	47	-3.29
R superior temporal gyrus	39	17	-26	45	-3.30
R inferior frontal gyrus	30	26	-20	59	-3.34
L superior parietal lobule	-27	-55	46	135	-3.58
L superior frontal gyrus	-9	11	55	62	-3.59
R middle frontal gyrus	48	50	-14	117	-3.94
L middle frontal gyrus	-54	26	34	936	-4.49
R middle frontal gyrus	54	35	28	531	-5.41
Occipital poles	-15	-100	-5	2539	-10.98

SI4. Whole-brain analysis identifying regions associated with pre- to post-intervention bias change

Neural regions associated with attitude change

Exploratory whole-brain analyses of the intervention tasks identified areas associated with later decreases in implicit bias from T1 to T3. No clusters survived FDR correction. We ran an additional analysis to determine the specificity of our ROI results by exploring whether additional regions beyond those hypothesized showed effects at a more liberal threshold; Figure 2 shows clusters at $p<.005$, $k=10$. Increased activity within the bilateral TPJ, bilateral fusiform gyri near the fusiform face area (FFA), and substantia nigra/ventral tegmental area extending to thalamus ($p<.005$, $k=10$) during the priming tasks was associated with later implicit attitude change (See Table SI4 for a complete table of whole-brain results).

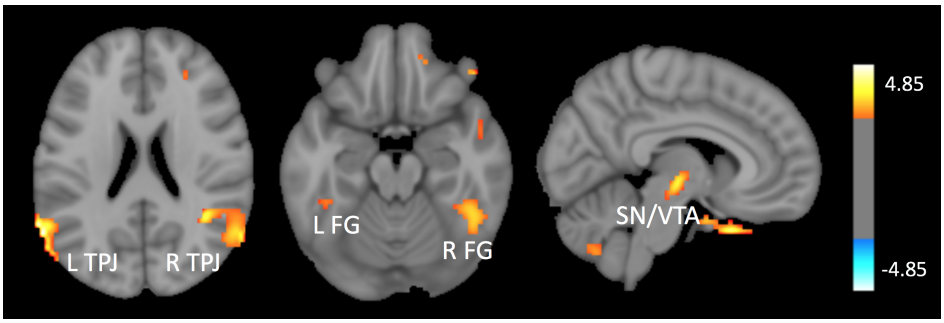

Figure SI4. *Whole-brain results showing active regions during the priming tasks that were associated with later decreases in implicit bias against drug addicts, p<.005, k=10. Note: L = left; R = right; TPJ = temporoparietal junction; FG = fusiform gyrus; SN/VTA = area near substantia nigra/ventral tegmental area.*

Table SI4. Whole-brain results during the intervention tasks showing regions associated with pre- to post-intervention decreases in bias against drug addicts, *p*<.005, *k*=10. Note: L = left; R = right

Region	*x*	*y*	*z*	*size*	*t*
L temporoparietal junction	-63	-58	25	172	4.99
orbitofrontal cortex	3	17	-35	105	4.60
R temporoparietal junction	39	-49	19	251	3.98
L posterior cerebellum	-36	-70	-50	106	3.94
R middle frontal gyrus	30	35	-14	77	3.81
R fusiform gyrus	45	-46	-17	205	3.73
substantia nigra/ventral tegmental area	9	-16	-5	36	3.72
L inferior frontal gyrus	-39	29	13	23	3.61
R interior frontal gyrus	36	5	31	54	3.53
L anterior cerebellum	-24	-46	-38	53	3.43
R fusiform gyrus	36	-7	-32	79	3.38
L superior frontal gyrus	-21	26	49	58	3.31
R middle frontal gyrus	30	32	49	77	3.16
R inferior frontal gyrus	48	41	16	12	3.09
R middle temporal gyrus	51	5	-23	30	3.05
R thalamus	24	-25	1	13	3.02
L middle frontal gyrus	-36	38	-8	12	3.02
L posterior cerebellum	-6	-73	-38	14	2.95
ventral striatum	18	17	-2	14	2.93
R cerebellum	9	-67	-47	19	2.92
L fusiform gyrus	-33	-10	-29	14	2.81