Neural Indicators of Anhedonia: Predictors and Mechanisms of Treatment Change in a Randomized Clinical Trial in Early Childhood Depression

Supplemental Information

Supplemental Methods and Materials

Recruitment

Young children (aged 3.0-6.11) from the St. Louis metropolitan area were screened and recruited from preschools, daycares, primary care, and mental health facilities. Inclusion criteria were: 1) meeting early onset major depressive disorder (MDD) symptom criteria on the K-SADS-early childhood (see below), with the validated syndrome requiring 4 instead of 5 symptoms of MDD; 2) no autism spectrum disorder; 3) no serious neurological syndrome or chronic medical disorder; 4) no significant developmental delay; and 5) no antidepressant medication or ongoing psychotherapy.

We obtained N=1378 Preschool Feelings Checklists (PFC), a validated brief screening measure with good sensitivity and specificity for early childhood depression. [1] Those with a score $\geq$ 3 (N=811) had a more extensive phone screen. This phone screen included the Preschool Age Psychiatric Assessment (PAPA) Major Depressive Disorder (MDD) module used to assess for potential inclusion and exclusion criteria. If a child met criteria for early onset MDD symptoms on the PAPA (the validated syndrome which requires 4 instead of 5 symptoms of MDD) and did not have an Autism Spectrum Disorder, a serious neurological or chronic medical disorder, or a significant developmental delay, then they were invited for an in-person assessment (N=369). Children were excluded if they were on antidepressant medications or in ongoing psychotherapy, but were not excluded if they were on stable doses of other psychotropic medications without antidepressant properties (e.g. Guanfacine, stimulants). If a child was too severely depressed to wait 18 weeks for treatment, as indexed by the child/family in serious acute distress, they were

referred for immediate treatment so as to preclude them being randomized to waitlist. All children who met these inclusion criteria then participated in a comprehensive mental health and emotional development baseline assessment at the WUSM EEDP. Children who met full criteria for early childhood MDD during the in person assessment were randomized to PCIT-ED or WL, with randomization stratified by gender and comorbid externalizing disorders.

ERP Tasks

Doors Guessing Task

Prior to the task, the experimenter first showed children three containers of prizes, each increasing in attractiveness to the child and in amount of 'points' required to obtain a prize. The experimenter told the children that if they received a certain number of points in the subsequent task, they could receive a prize from one of the containers. This exchange was designed to encourage the child to engage in the task and to make it relevant to the children. The order and timing of all stimuli were as follows (see Figure S2): (i) the text "Click for the next round" was presented until the participant pressed a button, (ii) a fixation cross was presented for 1000 ms, (iii) the graphic of two doors was presented until a choice was made, (iv) a fixation cross was presented for 1000 ms, (v) a feedback arrow was presented for 2000 ms, and finally (vi) a fixation cross was presented for 1500 ms. Participants responded using a Logitech Gamepad F310 game controller by pressing a specific button on the left of the controller with their left hand to choose the left door, or a specific button on the right with their right hand to choose the right door. A green upward arrow indicated a correct guess and a red downward arrow indicated an incorrect guess. All cues and feedback were presented against a black background and occupied approximately 3° of the visual field vertically and 1° horizontally. Participants were told that they would gain 10 points each time they opened a correct door and lose 5 points each time they opened an incorrect door. They were told that the experimenter would keep track of the points for them. Participants received negative feedback on exactly 50% of the trials, and positive feedback on exactly 50% of

the trials.

Picture Task

Prior to each picture, a fixation cross was presented on the screen for 1500 ms (see Figure S3). Pictures were than displayed for 1000 ms, followed by a left or right arrow (i.e., < or >) that appeared on the screen for 500 ms or until children responded with the controller. Children were placed approximately 60 cm from the screen and each picture occupied approximately 40 degrees of visual angle horizontally and vertically. Subjects first viewed a practice series of 10 pictures to familiarize them with the task procedure. After the practice children performed 80 trials – each picture was presented twice. The order of the trials was randomly determined for each child participant. Children had a break for up to 5 minutes after they completed 40 trials, that is at the halfway point in the task.

Psychophysiological Recording and Data Reduction

The electroencephalography (EEG) electrodes were attached while participants watched a movie of their choice. The EEG was recorded using a BrainVision ActiCHamp recording system and actiCAP active electrodes (Brain Products GmbH, Munich, Germany). The electrodes were mounted in an elastic cap using a subset of the International 10/20 System sites (FP1, F3, F7, FC1, FC5, FT9, C3, T7, CP1, CP5, TP9, P3, P7, O1, Fz, Cz, Pz, Oz, FP2, F4, F8, FC2, FC6, FT10, C4, T8, CP2, CP6, P4, P8, TP10, O2). A ground electrode was located at FPz. The EEG data were recorded and referenced to Cz. The horizontal electrooculogram (EOG) was recorded as the voltage between electrodes placed lateral to the external canthi and was used to measure horizontal eye movements. The vertical EOG was recorded from electrodes placed above and below the right eye and was used to detect blinks and vertical eye movements. An electrode on the forehead above the left eye served as the ground for the EOG signals. The EEG and EOG were digitized at 500 Hz with 24 bits of resolution.

Offline analysis was performed using Brain Vision Analyzer software (Brain Products GmbH, Munich, Germany). EEG data were re-referenced offline to the average of TP9 and TP10 (located adjacent to the mastoids) and band-pass filtered with half-power cutoffs at 0.1 and 30 Hz. The EEG was segmented for each trial, beginning 200 milliseconds before feedback onset (RewP) or picture onset (LPP) and continuing for 1,000 milliseconds. The EEG for each trial was corrected for blinks and eye movements using the Gratton et al method (11). Specific intervals for individual channels were rejected in each trial using an automated procedure, with physiological artifacts identified by the following criteria: a voltage step of more than 50.0 μV between sample points, a voltage difference of 300.0 μV within a trial, and a maximum voltage difference of less than 0.50 μV within 100-millisecond intervals.

Supplemental Results

Intent to Treat Analyses

As noted in the main text, there were some children who completed or had good ERPs at baseline, but not at post-treatment. The analyses in the main text focused on those children with usable ERPs both at baseline and post-treatment. To conduct intent-to treat analyses, we ran general linear models, separately for the Doors and Picture tasks, to impute post-treatment ERP scores for children with missing data, using baseline ERP scores, age, gender, and baseline PFC scores in the imputation. The results of these intent-to-treat analyses using general linear models were identical to those presented in the main text. Specifically, the Doors Task analysis again showed a significant effect of treatment group on the Win_{resid} score Post Assessment, after controlling for Age, PFC score, Win_{resid} at Baseline, and $Loss_{resid}$ at Post Assessment, t=2.35, df=471.58 p=.019. As in the main text, the picture task analysis of $Pleasant_{resid}$ at Post Treatment, with age, PFC scale-scores, and Baseline $Pleasant_{resid}$ as covariates again indicated no significant

treatment group differences in the residualized responses to pleasant pictures *t*=-1.50, df=174.90 *p*=.134.

Analyses Using Non-residualized Scores for Doors Guessing Task

Response to Treatment

A one-way ANCOVA was conducted to test for group differences between the treatment and waitlist groups in their response to win and the RewP (Win – Loss) at the Post-assessment. Covariates included age, PFC scale-scores, and baseline win, loss, or RewP amplitudes. There was a significant effect of group on the Win condition at the Post Assessment, $F_{(1,87)}$=5.48, p<.05, partial η^2=.06. The mean score for the treatment group (M=6.63, SD=0.98) was significantly more positive than the waitlist group (M=3.30, SD=1.00). There were no differences between the treatment and waitlist groups for the RewP $F_{(1,87)}$=2.26, p=0.14, partial η^2=.03.

Does Change in ERP Response to Win or Loss Predict Change in Depressive Symptoms?

Partial correlations controlling for age demonstrated that in the treatment group, a greater increase in Win from pre to post treatment was not associated with a greater decrease in MDD symptoms (r=-.17, p=.13), PFC score (r=.03, p=.41) or anhedonia (r=-.16, p=.15).

Linear regressions were conducted predicting remission from depression and/or change in depressive symptoms from the Win scores controlling for age. These analyses were conducted in the treatment group only. Baseline response to Win positively predicted the change in MDD core scores from baseline to follow-up (B=0.08; SE=0.04; t=2.09; p=0.04).

Analyses Using Non-residualized Scores for Picture Task

Response to Treatment

A one-way ANCOVA was conducted to test for group differences between the treatment and waitlist groups in their response to positive versus neutral pictures at the Post-Assessment. Covariates included age, PFC scale-scores, and baseline positive or neutral response amplitudes. There were no significant treatment group differences in positive or neutral responses at the Post-Assessment from 250-600ms (all *ps* > .58).

Does Change in LPP Response to Positive Pictures Predict Change in Depression?

Partial correlations controlling for age demonstrated that in the treatment group, a change in Positive from pre to post treatment was not associated with change in depressive symptoms (*r*=.07, *p*=.34), PFC scale-scores, (*r*=-.05, *p*=.37), or anhedonia (*r*=.22, *p*=.09).

Do Baseline LPP Responses to Positive Pictures Predict Treatment Outcome?

Linear regressions were conducted predicting remission from depression and/or change in depressive symptoms from the Positive pictures amplitudes controlling for age. These analyses were conducted in the treatment group only. Baseline positive amplitudes from 250-600ms predicted remission from depression (B=0.06; SE=2.51; *OR*=1.06; *p*=0.04). There was also a trend for Positive amplitude from 250-600ms to predict change in MDD symptoms from Baseline to Post Treatment (B= -0.05; SE=0.02; *t*= -1.92; *p*=0.06). However, positive amplitudes did not predict the change anhedonia scores (B= -0.01; SE= 0.004; *t*= -1.29; *p*=0.20).

Attention/Engagement Post PCIT Treatment

Examination of Figure 1 in the main text suggests that there may also be treatment effects on an P2/N2 complex in the Doors task, which may reflect enhanced engagement or attention in children in the PCIT-ED group compared to those in the WL group post treatment. To address this, we competed the same analyses that we had computed for the RewP, but focused on the P2/N2 complex noted by the reviewer. We used 200 ms to 300 ms post feedback so as not to

overlap with the time frame of the RewP. This analysis indicated no significant effects of treatment for either Pz ($F_{(1,87)}$=3.63, p=.06, partial η^2=.044) or Cz ($F_{(1,87)}$=1.21, p=.27, partial η^2=.015), though admittedly the effect is close for Pz. When we added this 200 to 300 ms component to the model for the RewP presented in the main text, the effect for 300-500 ms remained marginally significant (p=.098, partial η^2=.032). However, we did not see a similar P2/N2 effect in the picture task for pleasant pictures (see Figure S8). Further, while we did not have behavioral data from doors task that we could examine, we did examine the number of times children pressed the button for the positive picture pre and post treatment as a measure of attention/engagement. We found a robust main effect of pre-post treatment, such that all children responded (i.e., pushed the button) more frequently post-treatment than pre-treatment for pleasant pictures (p=.006, η^2=.109), but there was no interaction with treatment group (p = .575, η^2=.05). Thus, we would argue that while there may be some evidence for enhanced engagement post PCIT-ED in the doors task, this seems to be about reward feedback more specifically, and not salient stimuli in general, which would be consistent with enhanced reward processing.

Table S1. Demographic Characteristics of Participants Completing vs. Not Completing an ERP

	ERP Completed (n= 124)	**ERP Not Completed (n= 32)**	Group Comparison
Sex (% male)	65	72	X^2_1= 0.49, p=0.48
Race			X^2_2= 1.65, p=0.44
% White	78	72	
% African American	11	19	
% Other	11	9	
Age y, mean (SD)	5.60 (0.85)	5.32 (0.94)	t_{154}= -1.62, p=0.11
Preschool Feelings Checklist (PFC) scale score, mean (SD)	39.92 (10.75)	42.03 (10.90)	t_{154}= 0.99, p=0.33
Anhedonia sum score	1.37(1.06)	1.31(1.18)	t_{154} = -0.27, p=0.79
% Co-morbid externalizing disorders	53	53	X^2_1= 0.00, p=0.99
% on non-antidepressant medications	5	3	X^2_1= 0.17, p=0.68
Income-to-Needs, mean (SD)	2.92 (1.32)	2.65 (1.19)	t_{154}= -1.01, p=0.31

Abbreviations: SD = standard deviation; % = percent; t=t-test test statistic; p=p-value; X^2=Chi-Square test statistic

Table S2. Demographic Characteristics of Participants with Usable vs. Unusable ERP Data

	Reward Positivity Analyses			Late Positive Potential Analyses		
	Usable (N = 118)	Not Usable (N = 6)	Group Comparison	Usable (N = 99)	Not Usable (N = 25)	Group Comparison
Sex (% male)	64	83	X^2_1= 0.90, p=0.34	66	64	X^2_1= 0.03, p=0.87
Race						
% White	79	67	X^2_2= 0.50, p=0.78	79	76	X^2_2= 0.10, p=0.95
% African American	10	17		10	12	
% Other	11	17		11	2	
Age y, mean (SD)	5.64 (0.84)	4.79 (0.54)	t_{122}= -2.45, p = 0.02	5.73 (0.82)	5.01 (0.75)	t_{122}= -3.54, p=0.001
Preschool Feelings Checklist (PFC) scale score, mean (SD)	39.71 (10.70)	44 (11.97)	t_{122}= 0.95, p=0.34	39.94 (11.12)	39.84 (9.37)	t_{122}= -0.04, p=0.97
Anhedonia sum score	1.39(1.05)	1.00(1.27)	t_{122}= -0.88, p=0.38	1.39(1.09)	1.28 (0.98)	t_{122}= -0.48, p=0.63
% Co-morbid externalizing disorders	51	100	X^2_1= 5.54, p=0.02	54	52	X^2_1= 0.02, p=0.89
% on non-antidepressant medications	5	17	X^2_1= 1.92, p=0.17	2	16	X^2_1= 8.47, p=0.004
Income-to-Needs, mean (SD)	2.92 (1.32)	2.88 (1.57)	t_{122}= -0.06, p=0.95	2.94 (1.30)	2.81 (1.44)	t_{122}= -0.43, p=0.67

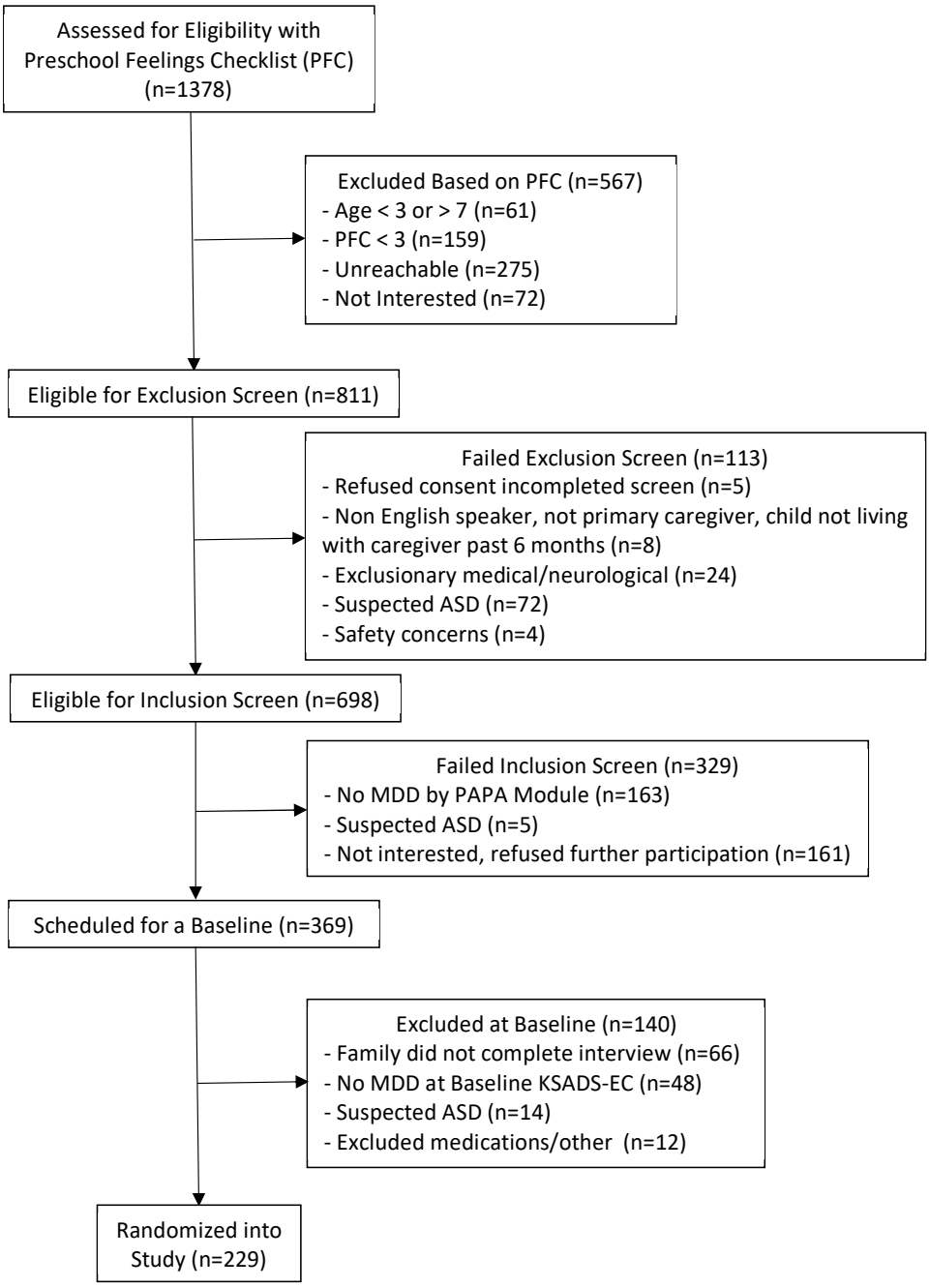
Abbreviations: SD = standard deviation; % = percent; t=t-test test statistic; p=p-value; X^2=Chi-Square test statistic

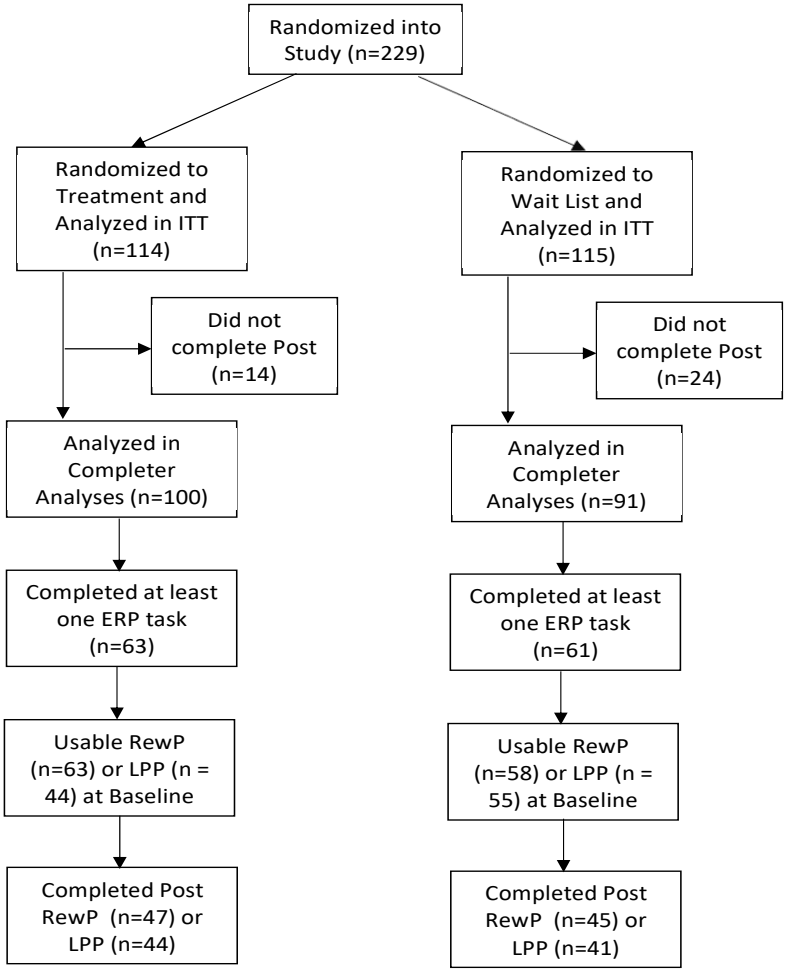
Table S3. Post Assessment Severity and Diagnostic Characteristics in PCIT-ED and Wait List Subjects with ERP Data

	Wait List		PCIT-ED		Wait List vs. PCIT-ED				
Primary Outcome: MDD Diagnosis	**%**	**N**	**%**	**N**	χ^2	**p**	**FDR p**	**OR**	**95% CI**
Major Depressive Disorder (or NOS)	76.2	32	21.4	9	21.38	<0.0001	<0.0001	12.00	(4.19, 34.42)
Secondary Outcome: Remission	**%**	**N**	**%**	**N**	χ^2	**p**	**FDR p**	**OR**	**95% CI**
Remission of MDD*	23.8	10	73.8	31	18.34	<0.0001	<0.0001	0.11	(0.04, 0.31)
PFC-Scale reduced ≥50%, no MDD	4.8	2	45.2	19	12.18	0.0005	0.0007	0.06	(0.01, 0.30)
Secondary Outcome: Severity	**Mean**	**SD**	**Mean**	**SD**	**t**	**p**	**FDR p**	**Partial η^2**	**Cohen's d**
MDD core score	4.24	2.02	1.62	1.61	6.24	<0.0001	<0.0001	0.33	1.05
PFC-Scale score	31.98	12.29	18.86	8.41	5.71	<0.0001	<0.0001	0.29	1.06
Secondary Outcome: Impairment	**Mean**	**SD**	**Mean**	**SD**	**t**	**p**	**FDR p**	**Partial η^2**	**Cohen's d**
CGAS score	55.60	17.66	78.38	16.79	-5.65	<0.0001	<0.0001	0.29	1.27
PECFAS/CAFAS	7.58	3.51	4.55	3.11	4.13	<0.0001	<0.0001	0.18	0.81
CGI-I score	3.40	1.21	2.07	0.87	5.63	<0.0001	<0.0001	0.28	1.28

Cohen's d is for the change from baseline to post; FDR = false discovery rate; OR = odds ratio; *Remission defined as not meeting diagnostic criteria for major depressive disorder and a 50% or greater reduction in MDD core score from baseline to post; Analyses covary for baseline characteristics, gender, and baseline externalizing disorder.

Figure S1: Consort Diagram of ERP Substudy of the PCIT-ED Study[1]





1. Luby JL, Barch DM, Whalen D, Tillman R, Freedland KE. A Randomized Controlled Trial of Parent-Child Psychotherapy Targeting Emotion Development for Early Childhood Depression. *American Journal of Psychiatry.* in press.

Figure S1. Consort Diagram for Study

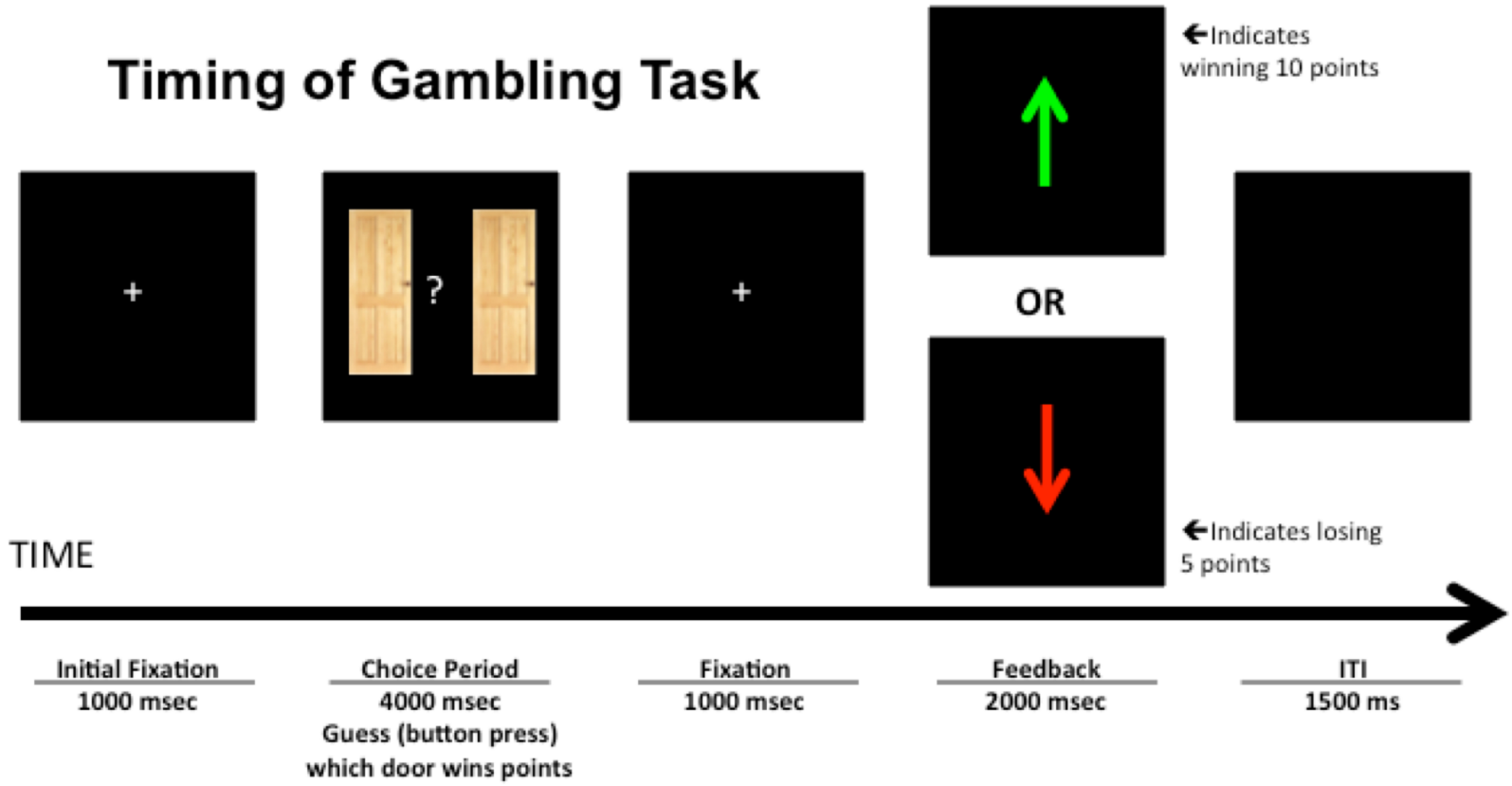


Figure S2. Illustration of the Doors Guessing Task

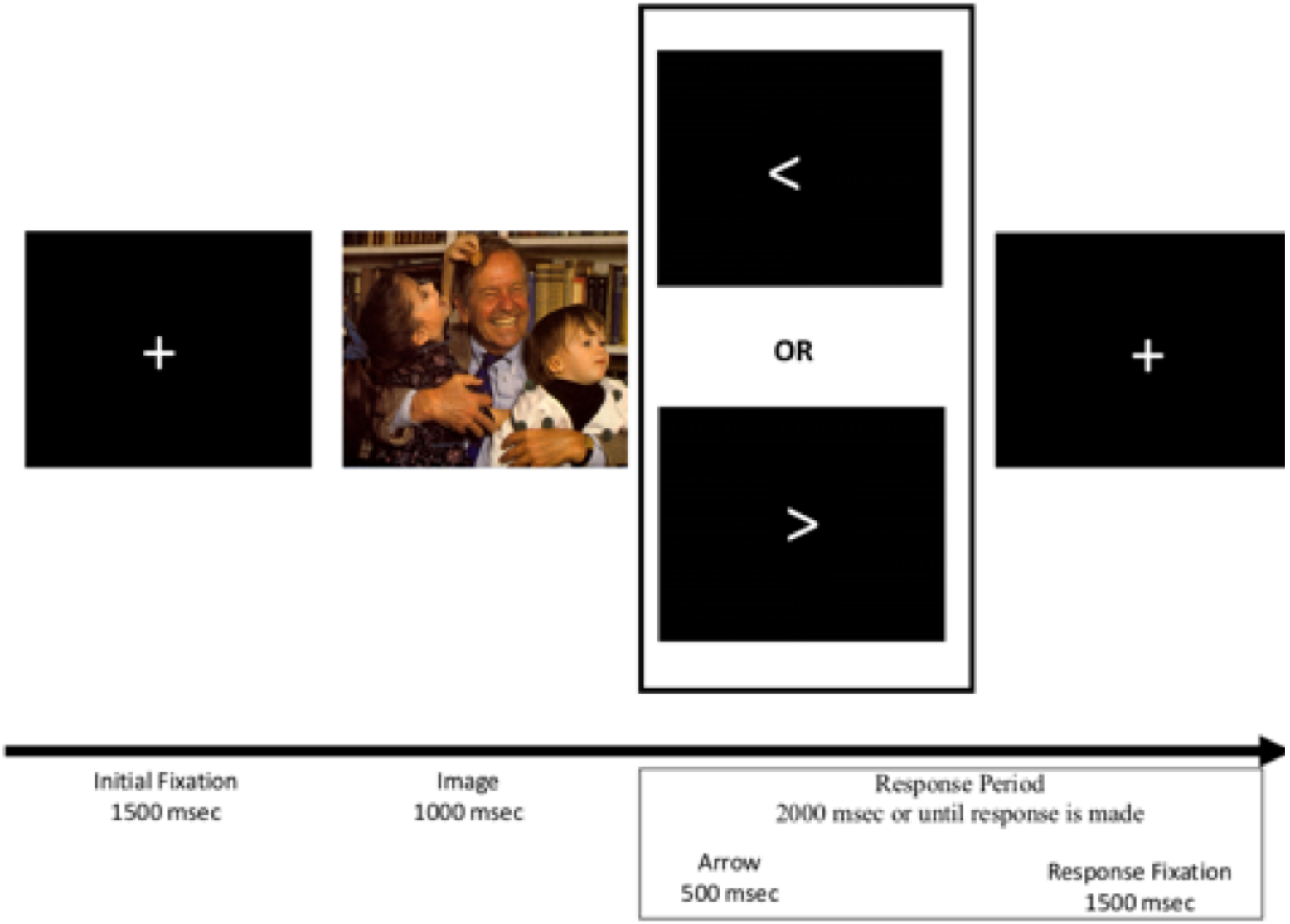


Figure S3. Illustration of the Positive and Negative Picture Task

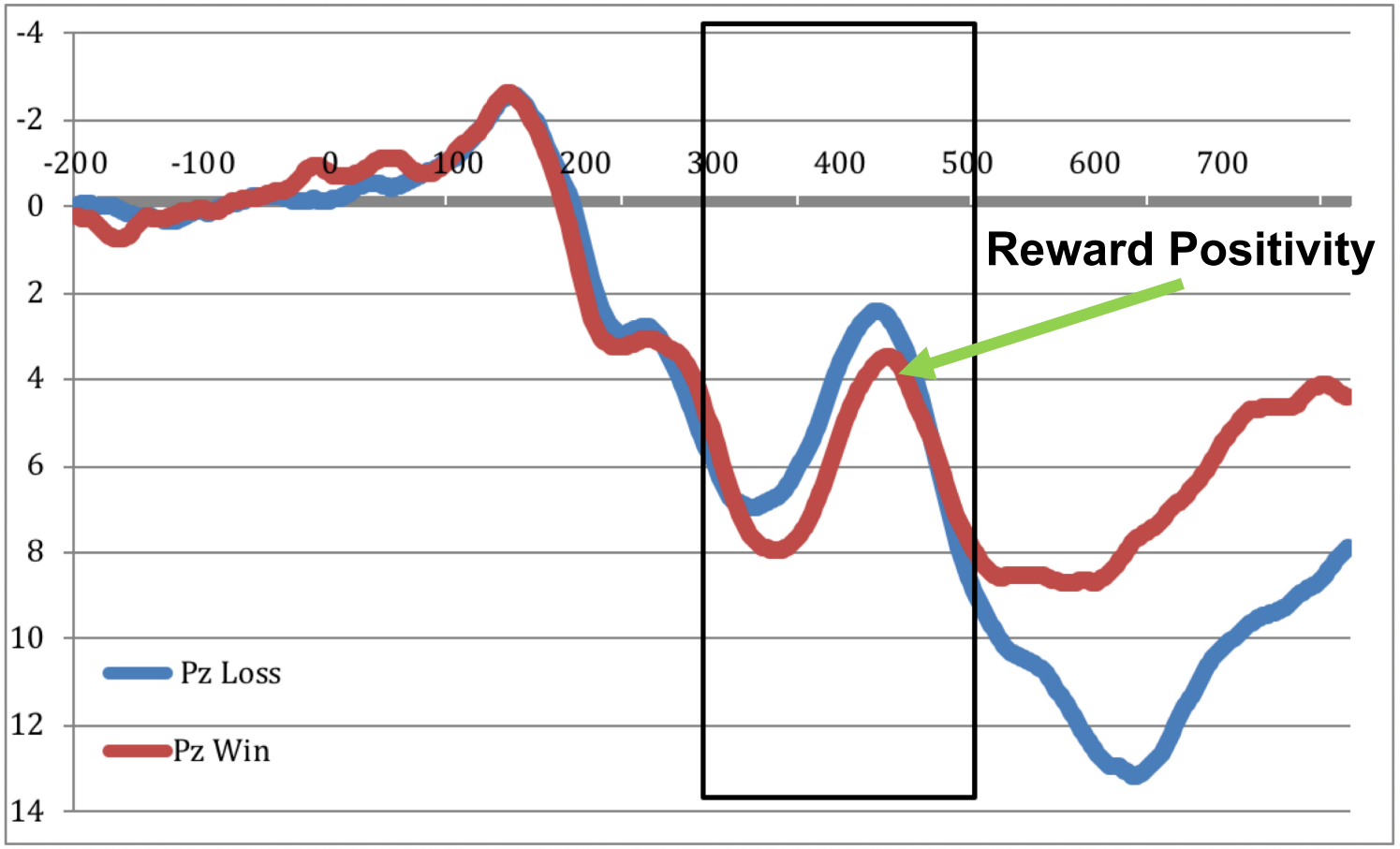


Figure S4. Reward Positivity Grand Averages: Grand average ERP waveforms at Pz collapsed across all children and all data from Baseline and Post Treatment Assessment for both the Reward and Loss conditions. The red line is responses to wins and the blue line is responses to losses. Voltages are plotted with the more negative values at the top of the graph, as is the frequent convention in ERP reports.

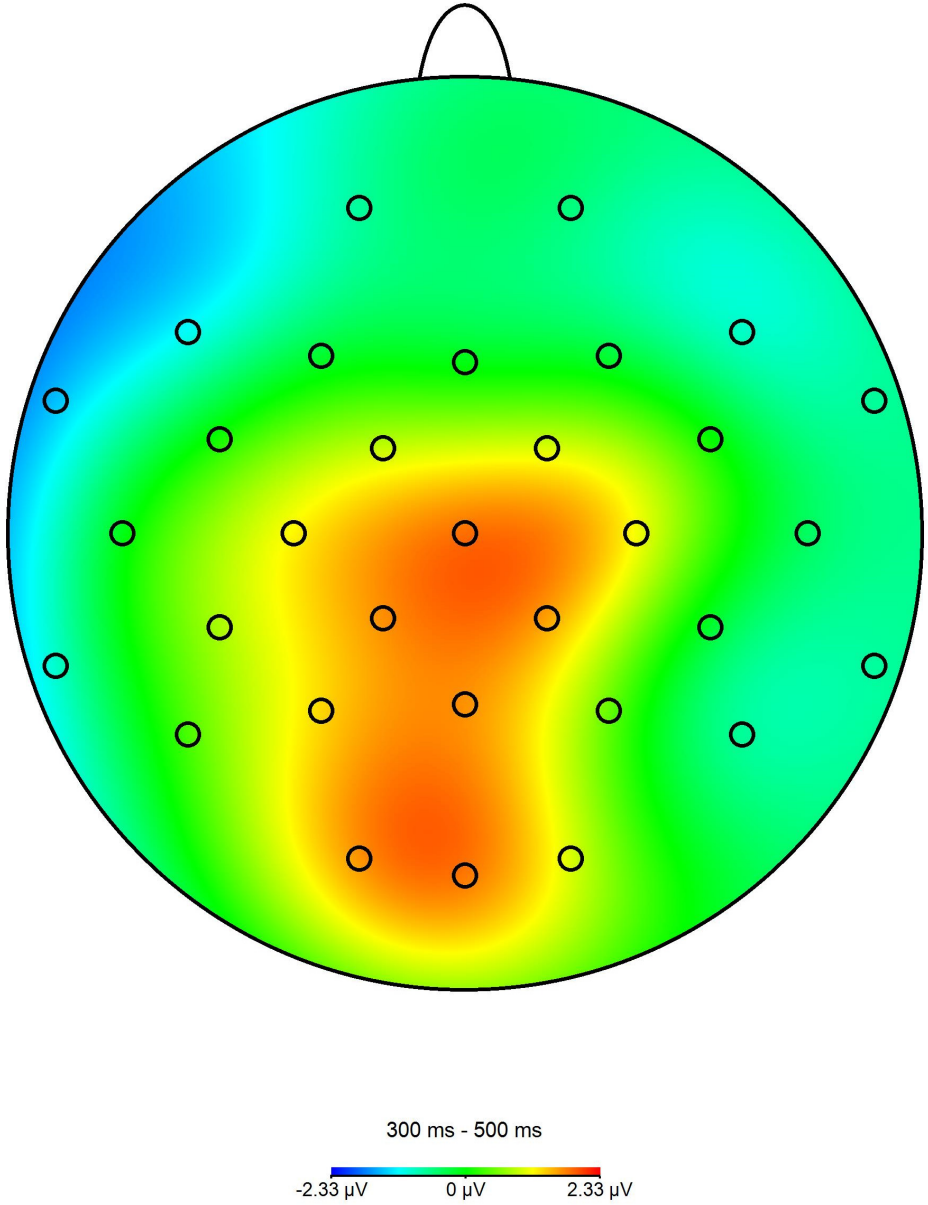


Figure S5. Reward Positivity Head Map: Distribution of grand average ERPs from 300 ms to 500ms in the Win condition of the Doors Task collapsed across all children and all data from Baseline and Post Treatment Assessment for both the Reward and Loss conditions. Voltages are plotted with the more positive values in warm colors and more negative values in cool colors.

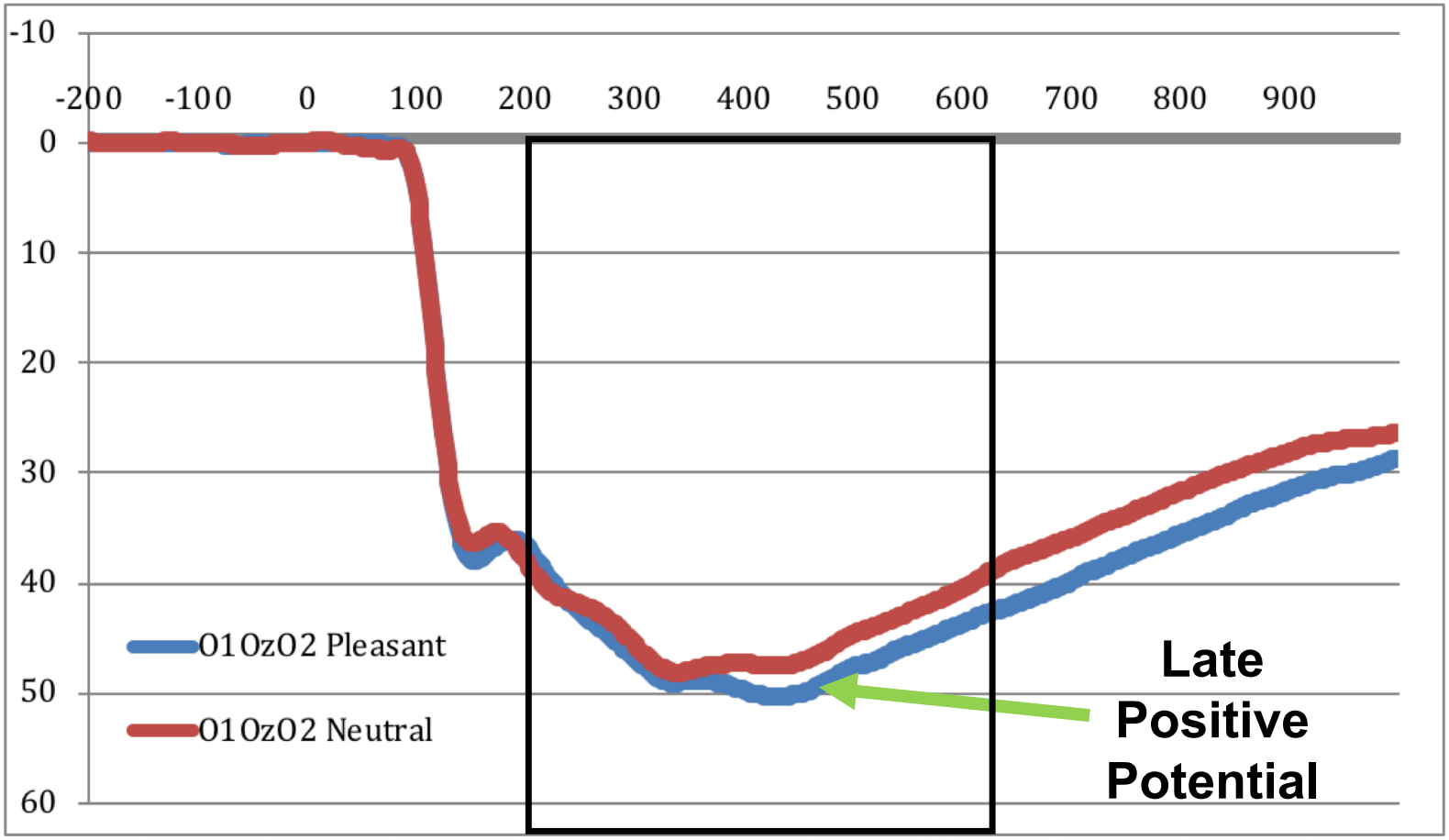


Figure S6. Late Positive Potential Grand Averages: Grand average waveforms from the average of O1, Oz, and O2 across all children and all data from Baseline and Post Treatment Assessment for both the pleasant and neutral picture conditions. The blue line is responses to pleasant pictures and the red line is responses to neutral pictures. Voltages are plotted with the more negative values at the top of the graph, as is the frequent convention in ERP reports.

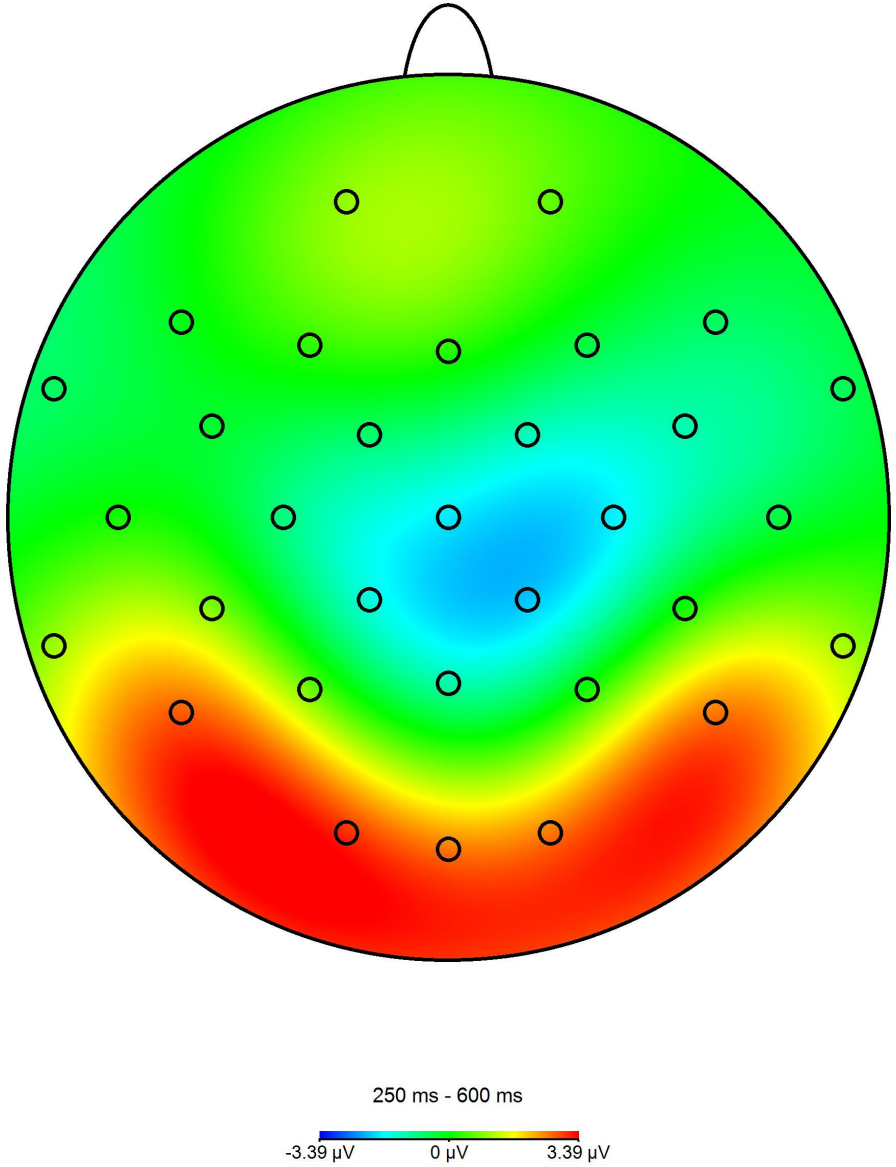


Figure S7. Late Positive Potential Positivity Head Map: Distribution of grand average ERPs in the Positive Picture Condition from 250 ms to 600 ms collapsed across all children and all data from Baseline and Post Treatment Assessment for both the Reward and Loss conditions. Voltages are plotted with the more positive values in warm colors and more negative values in cool colors.

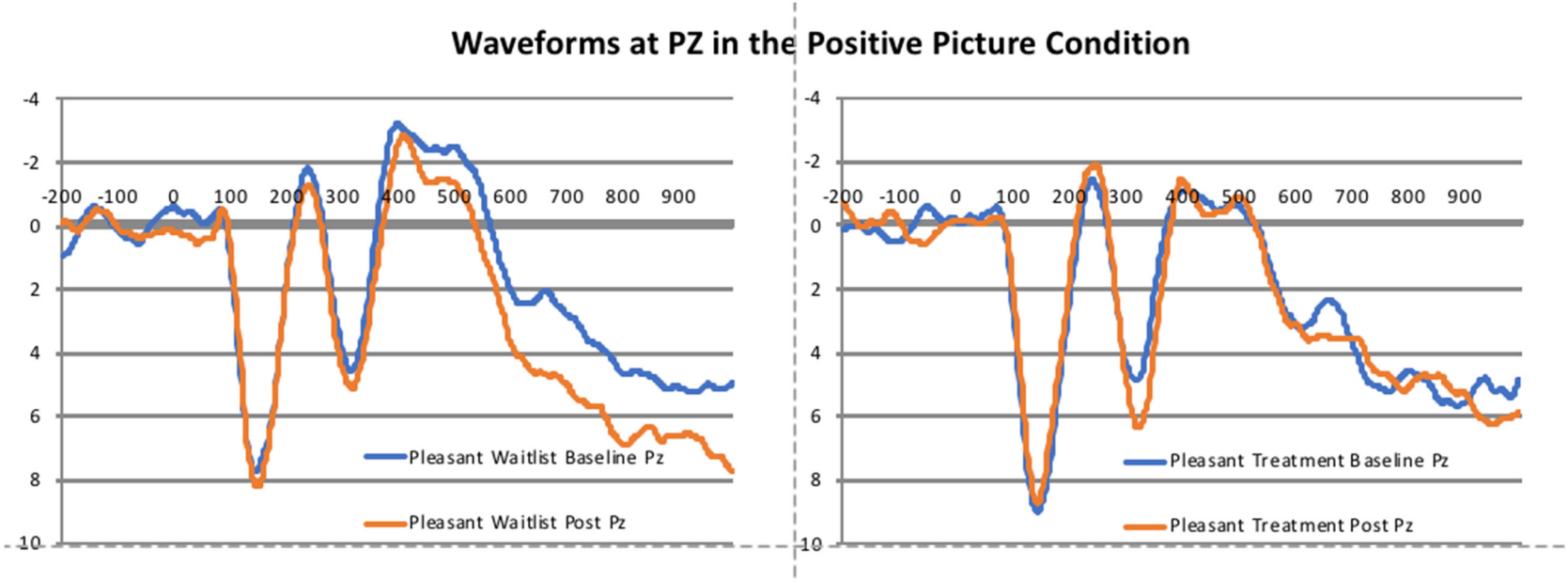


Figure S8. Late Positive Potential Grand Averages at PZ*:* Grand average waveforms for PZ across all children and all data from Baseline and Post Treatment Assessment for the pleasant picture condition. Voltages are plotted with the more negative values at the top of the graph, as is the frequent convention in ERP reports.

Supplemental References

1. Luby JL, Heffelfinger A, Koenig-McNaught AL, Brown K, Spitznagel E. ThePreschool Feelings Checklist: a brief and sensitive screening measure for depression in young children. *Journal of the American Academy of Child and Adolescent Psychiatry.* 2004;43(6):708-717.

2. Eyberg SM, Funderburk BW, Hembree-Kigin TL, McNeil CB, Querido JG, Hood KK. Parent-Child Interaction Therapy with behavior problem children: One and two year maintenance of treatment effects in the family. *Child & Family Behavior Therapy.* 2001;23(4):1-20.

3. Saarni C. *The Development of Emotional Competence.* New York, NY: Guilford Press; 1999.

4. Lenze SN, Pautsch J, Luby J. Parent-child interaction therapy emotion development: a novel treatment for depression in preschool children. *Depress Anxiety.* 2011;28(2):153-159.

5. Bress JN, Meyer A, Hajcak G. Differentiating anxiety and depression in children and adolescents: evidence from event-related brain potentials. *J Clin Child Adolesc Psychol.* 2015;44(2):238-249.

6. Bress JN, Smith E, Foti D, Klein DN, Hajcak G. Neural response to reward and depressive symptoms in late childhood to early adolescence. *Biol Psychol.* 2012;89(1):156-162.

7. Foti D, Kotov R, Klein DN, Hajcak G. Abnormal neural sensitivity to monetary gains versus losses among adolescents at risk for depression. *Journal of abnormal child psychology.* 2011;39(7):913-924.

8. Foti D, Carlson JM, Sauder CL, Proudfit GH. Reward dysfunction in major depression: multimodal neuroimaging evidence for refining the melancholic phenotype. *Neuroimage.* 2014;101:50-58.

9. Nelson BD, Perlman G, Hajcak G, Klein DN, Kotov R. Familial risk for distress and fear disorders and emotional reactivity in adolescence: an event-related potential investigation. *Psychol Med.* 2015;45(12):2545-2556.

10. Lang PJ, Bradley MM, Cuthbert BN. *International affective picture system (IAPS): Technical manual and affective ratings.* Gainesville: University of Florida;1999.

11. Gratton G, Coles MG, Donchin E. A new method for off-line removal of ocular artifact. *Electroencephalography and clinical neurophysiology.* 1983;55(4):468-484.