Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

HHS Public Access

Author manuscript

Biol Psychiatry Cogn Neurosci Neuroimaging. Author manuscript; available in PMC 2018 April 01.

Published in final edited form as:

Biol Psychiatry Cogn Neurosci Neuroimaging. 2017 April ; 2(3): 289–297. doi:10.1016/j.bpsc.2016.12.004.

Internal Consistency of Functional Magnetic Resonance Imaging and Electroencephalography Measures of Reward in Late Childhood and Early Adolescence

Dr. Katherine R Luking, PhD,
Psychology Department at Stony Brook University

Dr. Brady D Nelson, PhD,
Psychology Department at Stony Brook University

Dr. Zachary P Infantolino, PhD,
Psychology Department at Stony Brook University

Dr. Colin L Sauder, PhD, and
Psychiatry Department at the University of Texas Health Science Center at San Antonio

Dr. Greg Hajcak, PhD
Psychology Department at Stony Brook University

Abstract

Background—Abnormal neural response to reward is increasingly thought to function as a biological correlate of emerging psychopathology during adolescence. However, this view assumes such responses have good psychometric properties—especially internal consistency—an assumption that is rarely tested.

Methods—Internal consistency (i.e., spilt-half reliability) was calculated for event-related potentials (ERPs) and Blood Oxygen Level Dependent (BOLD) responses to monetary gain and loss feedback from the same sample of 8–14 year-old females (n=177). Internal consistency for ERPs (i.e. feedback negativity) and BOLD responses within the ventral striatum and medial/lateral prefrontal cortex to gain, loss, difference scores (gain-loss), and residual scores (gain controlling for loss) were compared. Moderation analyses were conducted to investigate whether internal consistency differed by age.

Results—ERP and BOLD responses to gain and loss feedback showed high internal consistency in all regions (*Spearman Brown Coefficients* (*SB*) ≥ 0.70). When considering difference and residual scores, however, responses showed lower internal consistency (*SB*s ≤ 0.50), with

Correspondence: Katherine R. Luking, Department of Psychology, Stony Brook University, Stony Brook, NY, 11794, katherine.luking@gmail.com, Phone: 502-445-0921.

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.

DISCLOSURES
All authors report no biomedical financial interests or potential conflicts of interest.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

particularly low internal consistency for subtraction-based scores ($SB \leq 0.36$). Age was not a significant moderator of split-half relationships, indicating similar internal consistency across late childhood to early adolescence.

Conclusions—Within the same subjects, high internal consistency was observed for both ERP and fMRI measures of response to gains and losses, which did not vary as a function of age. Moreover, excellent psychometric properties were evident even within the first half of the experiment. Difference scores were characterized by lower internal consistency, although regression-based approaches outperformed subtraction-based difference scores.

Keywords

Reward; Internal consistency; Childhood; Adolescence; fMRI; ERP

INTRODUCTION

In the past decade efforts to leverage neuroscience-based measures to shed light on psychopathology have increased. For example, the National Institute of Mental Health's (NIMH) Research Domain Criteria (RDoC) initiative aims to move psychiatric research towards investigating neurobiological mechanisms of specific domains of function (1). Such biologically-oriented approaches to understanding psychopathology and individual differences require neural measures with good psychometric properties (2). No measure can be valid if it is not reliable (3, 4). Yet, few studies have considered measurement properties of neural metrics. In the current article we focus on internal consistency reliability (referred to as internal consistency throughout) of the neural response to reward feedback (both reward gains and losses) given its central role in the RDoC matrix's 'positive valence systems', and that abnormal response to reward is a promising biomarker of risk for multiple types of psychopathology (5, 6).

Establishing whether neural response to reward feedback shows adequate internal consistency is a critical first step in determining the utility of such individual difference measures. Neural responses to gains and losses show dramatic changes over the course of adolescence (7–9) with abnormal responses to reward prospectively predicting increases in depressive symptoms and substance use during adolescence (5, 6, 10, 11). Such findings have led to the idea that abnormal neural response to reward may confer vulnerability to depression and other disorders during development, when patterns of normative reward response are in flux (7). As such, investigating internal consistency of neural reward-response during emerging adolescence, a time of increasing onset of reward-related psychopathology particularly for females (12), is the aim of the current study.

Several studies have examined test-retest reliability of neural measures of reward response (e.g. (13–18)). Test-retest reliability refers to the consistency of individuals' scores assessed at different points in time. To assess test-retest reliability of blood oxygen level dependent (BOLD) responses and event related potentials (ERPs), participants would complete an identical task in two sessions separated by weeks or months. Neural responses are then correlated across sessions. High test-retest reliability indicates relative stability of scores over time—and suggests that a measure has trait-like properties. In the domain of reward,

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

ventral striatal response to reward feedback has been found to have relatively low test-retest reliability in some studies (13), whereas others report moderate test-retest reliability (14, 15). ERPs indexing reward response, including feedback negativity (FN) and reward positivity (RewP) (19, 20), tend to show higher test-retest reliability (16–18)—although these measures have never been directly compared in the same subjects. It is important to note that test-retest reliability is sensitive to between-session variations in methodological noise (e.g. thermal and system noise and motion artifacts in fMRI) and participant state (e.g. time of day, fatigue, or stress level) (21). Test-retest reliability does not address whether a measure is *internally* consistent. It is possible for a measure to be internally consistent despite low test-retest reliability (i.e., if a measure was highly sensitive to state-related variables); moreover, it is possible for a measure to be reliable over multiple administrations, despite having poor internal consistency (i.e., if a measure was comprised of multiple trait-like items that were themselves unrelated to one another).

Internal consistency refers to the similarity of items on a measure. A relatively simple method of calculating internal consistency is the split-half method, which involves correlating two separate scores for each subject that are derived from the same measure at the same testing session. Most frequently, scores on odd and even trials are utilized. Because splitting the data artificially reduces the number of trials by half, Pearson coefficients (r) are corrected using the Spearman–Brown prediction formula ($SB = 2r/(1+r)$).

Internal consistency is a fundamental measurement property that places an upper limit on how well a measure can index individual differences. A measure cannot correlate better with another measure than it correlates with itself; thus, poor internal consistency will limit the ability of a measure to relate to other individual difference variables. This point has been central to several debates regarding the reproducibility and implications of studies linking BOLD response to individual differences (21–25). However, the internal consistency of the BOLD response to reward has not been examined to date; in fact, internal consistency of fMRI measures is hardly ever examined and reported in fMRI studies. This fact is surprising given the increasing focus on how ventral striatal reward response relates to individual differences in symptom severity and risk factors for psychopathology both in adults (e.g. (26, 27)) and during adolescence (e.g. (28–32)). Reward-related ERPs, however, are characterized by high split-half internal consistency even during late childhood and early adolescence (17, 18, 33), but it is unclear whether internal consistency *changes* as a function of age during adolescence. Moreover, no studies have directly compared psychometric properties of ERP and fMRI measures of reward in the same subjects.

The current study aimed to assess internal consistency of both ERP and BOLD responses to reward gain and loss feedback in a large unselected sample of girls during emerging adolescence. Further, we test the hypothesis that internal consistency will be comparable across age during emerging adolescence despite the normative increases in response to reward feedback observed during these ages (e.g. (8)). This is critical not only for determining the utility of ventral striatal/ERP response to reward as an early biomarker of risk, but also for interpreting developmental changes in reward-related responses. That is, presumed “developmental” differences in reward-related neural responses could be attributed to variation in internal consistency as a function of age. We also compare the

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

internal consistency of BOLD and ERP responses to reward within the same subjects to examine potential methodological differences in assessing reward-related neural activity. Finally, as internal consistency may differ as a function of task length, particularly when a limited number of stimuli are repeated many times over a longer task, we report internal consistency of early and late components of the task. Investigating these properties of internal consistency during emerging adolescence will provide a critical baseline for future inquiries regarding relationships between neural reward response and psychopathology risk during this time of increasing vulnerability to psychopathology.

MATERIAL AND METHODS

Participants

198 girls aged 8–14, and their parents, from a parent study (described in (34) participated in an fMRI scan. This study focused on neural response to reward in females during emerging adolescence as studies in the depression-risk literature are increasingly focusing on females given the increased incidence of depression in women, a sex difference that emerges over adolescence (35). 177 girls provided sufficient quality data from both fMRI and EEG versions of the doors guessing task and are included in the current study (exclusions for excessive motion n=5, scanner sequence or other mechanical error n=16). Girls were aged 8–14 years (M=12.56; SD = 1.80), and were 77.4% Caucasian, 16.9% African American, 2.3% Hispanic, 3.4% identified as ‘Other’. Informed assent and consent were obtained from the participant and their parent, respectively, prior to participation. The Stony Brook University Institution Review Board approved the research protocol.

Measures

Doors Task—The doors task was similar to versions used in previous studies (36). Participants are presented with two identical doors and select the left or right door via mouse click. They are told that they will gain $0.50 or lose $0.25 on each trial depending on whether the ‘correct’ door is selected. The order of gain and loss feedback events was predetermined such that all participants experienced 30 gain feedback events and 30 loss feedback events presented in a pseudo-random order.

Task components were identical in fMRI and EEG task versions; only the timing of some task components differed slightly. First, the image of the doors was presented (fMRI – 3000ms; EEG – until the participant made a selection). After stimulus offset, a fixation cross (+) was presented (fMRI – 600ms; EEG – 1000 ms), followed by feedback (fMRI – 1000 ms; EEG – 2000 ms). A green arrow pointing upward (↑) represented gain feedback, and a red arrow pointing downward (↓) represented loss feedback. In the EEG version a final fixation cross (1500 ms) was displayed followed by instructions to ‘click a button’ for the next round. In the fMRI version the post-feedback inter-trial interval (mean 3200 ms, min 1100 ms, max 11600 ms) was followed by fixation (600 ms) signaling the next trial.

fMRI Data Acquisition and Processing—FMRI data were acquired using a 3T Siemens Trio whole body scanner. T2-weighted whole-brain volumes with an EPI sequence sensitive to BOLD signal were acquired (TR = 2100 ms, TE = 22 ms, flip angle = 83°,

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

matrix dimensions = 96 × 96, FOV =224 × 224 mm, slices = 40, slice thickness = 3.5 mm, and gap = 0 mm). Statistical parametric mapping (SPM8; Wellcome Department of Cognitive Neurology, Institute of Neurology, London, United Kingdom) was used to perform standard preprocessing procedures with default parameters, including image realignment corrections for head movements, slice timing corrections for acquisition order, normalization to standard 2 × 2 × 2 mm Montreal Neurological Institute space, and spatial smoothing with a Gaussian full-width-at-half-maximum 8 mm filter. A block of fixation (21s) preceded and followed the Doors task for majority of participants (n=152; 260 whole brain volumes). The remaining participants (n=25) experienced the same order of trials during the Doors Task, but without the two fixation blocks (242 whole brain volumes). Results were qualitatively similar when data from subjects without the additional fixation were excluded. ArtRepair version 4 (37) was used to identify and repair (interpolate) frames where frame to frame motion exceeded 2mm (1 voxel) or global signal deviations > 2%. Participants were excluded if >20% of frames were interpolated (n=5).

Event-related fixed-effects general linear models (GLMs) were created for each participant. Door cue, gain feedback, and loss feedback were modeled separately along with a regressor for interpolated frames and regressors for each of the 6 rigid body motion parameters estimated during realignment; the implicit baseline was comprised of unmodeled fixation. Two GLMs were created for each participant. In the first GLM, even and odd trials of each event type (door, gain, loss) were modeled separately with 15 trials of each feedback type included in each split-half. In the second GLM, even/odd trials of each event type were modeled separately for the first and second half of the task. To ensure equal numbers of events in each split-half, the final gain/loss even/odd trial were all modeled as a separate event type in the second GLM. Thus, 7 trials of each feedback type were modeled in each of the four split-halves.

Contrasts were created to examine activation to gain feedback, loss feedback, and subtraction-based difference score (SCOREs = gain>loss) for each split-half. Reward-related regions of interest (ROIs) were identified using a one-sample t-test for SCOREs; FWE $p<0.01$, cluster size > 20. The mean response within resultant regions of interest (ROIs) for each contrast was extracted for individual participants using the MarsBar toolbox (38). Values were imported into SPSS 23.0 for use in split-half analyses and to create a residualized score using regression (unstandardized residuals or SCOREr=gain controlling for loss). Internal consistency of residual scores were examined in addition to those of subtraction-based difference scores (SCOREs) as residual scores are commonly used in the ERP literature and tend to show better psychometric properties compared to difference scores (39).

EEG Recording and Processing—EEG was recorded using an elastic cap with 34 electrode sites placed according to the 10/20 system. Electrooculogram (EOG) was recorded using two electrodes placed approximately 1 cm outside of the right and left eyes and two placed approximately 1 cm above and below the right eye. EEG and EOG were recorded using the Active Two BioSemi system (BioSemi, Amsterdam, Netherlands) and digitized with a sampling rate of 1024 Hz using a low-pass fifth order sinc filter with a half-power cutoff of 204.8 Hz.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

EEG data were analyzed using Brain Vision Analyzer (Brain Products, Gilching, Germany). Data were referenced offline to the average of left and right mastoids, band-pass filtered (0.1 to 30 Hz), and corrected for eye movement artifacts (40). Feedback-locked epochs were extracted from 200 ms before to 800 ms following stimulus onset. The 200 ms prestimulus interval was used as the baseline.

Feedback-locked ERPs were averaged separately for gains and losses, which were then scored as the mean amplitude from 250–350 ms following feedback at FCz, where the difference between gains and losses was maximal. The difference between gain and loss trials was also quantified in two ways; first, by subtracting loss from gain trials (i.e., subtraction-based scores, SCOREs); second, by saving the unstandardized residual scores when predicting gain from loss trials using regression (i.e., regression-based residual scores, SCOREr, see (39)). All ERPs were scored separately for even and odd trials for use in split-half internal consistency analyses. See Supplemental Tables S1–S3 for the pattern of ERP deflections in each split half, paired t-tests comparing ERPs from odd and even trials, and relationships between gain and loss ERPs, respectively. Figure 1A displays the ERP waveforms following gains, losses, and SCOREs (gain-loss) as well as the 3-D rendered scalp distribution of SCOREs.

Internal Consistency Analyses—Pearson correlations coefficients (r) comparing response to gain, loss, SCOREs, and SCOREr on odd and even trials were calculated for each ROI/ERP and then adjusted using the Spearman-Brown prediction formula ($SB = 2r/(1+r)$) to correct for the data loss from splitting the number of trials in half. Spearman-Brown coefficients above 0.8 are considered ‘good/excellent’, above 0.7 are ‘acceptable’, and below 0.6 are ‘poor’.

Moderation analyses, testing whether internal consistency varied by age, were conducted using the PROCESS v2.13.2 macro for SPSS (41). Age was entered as a continuous moderator of the relationship between even and odd responses for each ROI/ERP. Given that six sets of moderation analyses were conducted, one for each ROI and one for ERPs, moderation was considered significant when $p<0.008$ for the age interaction. Exploratory internal consistency and age moderation analyses were also conducted separately for the first and second half of the data using the same methods.

The cocor package (42), implemented in R, was used to compare split-half internal consistency across imaging modalities and halves of the task. Silver, Hittner, and May’s (2004) modification of Dunn and Clark’s z (1969) using a back transformed average Fisher’s Z procedure are reported.

RESULTS

BOLD Response to Feedback

Response to gain feedback was greater than response to loss feedback within ventral striatum (−13,7,−8; 13,7,−8), medial prefrontal cortex (6,63,3), and lateral prefrontal cortex (−43,54,3; 29,58,−5), MNI coordinates (Supplemental Figure S1). See Supplemental Tables S1–S3 for activation patterns across odd/even averages, comparison of odd/even trials, and

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

relationships between gain and loss response. Portions of the bilateral parietal cortex and cerebellum also showed differential response to gain versus loss, but are not discussed. The subtraction based score (SCOREs) for ERPs and BOLD signal within the right ventral striatum were positively correlated ($r(175)=0.15$, $p=0.046$), BOLD response in other regions and for other contrasts did not significantly relate to ERPs (all $p>0.10$).

Split-Half Reliabilities of ERP and BOLD Response to Feedback

Although ERPs to gains and losses showed high internal consistency, *SBs*>0.85, SCOREs and SCORErshowed poorer internal consistency, with the lowest internal consistency for SCOREs, *SB*≤0.36 (Table 1; Figure 1B–D; see Supplemental Table S4 for Intraclass Correlation Coefficients).

Across ROIs, BOLD responses showed patterns of internal consistency similar to ERPs. BOLD response to both gain and loss showed moderate to high internal consistency, *SB*>0.66, and both SCOREs and SCORErwere characterized by lower internal consistency, *SB*≤0.48. SCOREs consistently showed lower internal consistency than SCORErr (Table 1; Figure 2).

Age Moderation

Age did not significantly moderate internal consistency of BOLD or ERP responses to gains, losses, or SCOREs. However, after multiple-comparisons correction there was a trend towards increasing internal consistency of VS SCORErwith age (Table 2), but this should be interpreted with caution given the low internal consistency of VS SCOREr.

Internal Consistency of ERP versus Ventral Striatal BOLD Responses to Feedback

Although both ERP and fMRI measures showed high internal consistency, internal consistency of ERP responses to gain and loss were significantly higher than that for VS BOLD responses (gain $z=-1.97$, $p=0.025$; loss $z=-3.40$, $p<0.001$). Internal consistency of SCOREs and SCORErwere similar for VS and ERP responses (SCOREs $z=-0.11$, $p=0.458$; SCORErr $z=-0.25$, $p=0.403$).

Internal Consistency of ERP and BOLD Responses to Feedback in the First and Second Half of Trials

ERP and BOLD responses to gain and loss were characterized by high internal consistency, even during the first half of the task (Table 3; see Supplemental Table S5 for Intraclass Correlation Coefficients). VS and ERP responses continued to show high internal consistency in the second half of the task, although internal consistency of prefrontal responses weakened in the second half of the task, particularly for gains (Table 3). Again, subtraction and residual scores had worse internal consistency during both halves of the task (Table 3).

Split-half relationships in both the first and second half of the task were not significantly moderated by age for ERPs and BOLD response in most ROIs (Supplemental Table S6). However, internal consistency of responses to gain and differences scores within the left prefrontal cortex tended to decrease with age during the second half of the task.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

DISCUSSION

The current study is the first to demonstrate that both ERP and fMRI measures of response to reward gain and loss during a simple gambling task have high internal consistency in the same female subjects during late childhood and early adolescence. These findings add to a growing literature reporting similar internal consistency for ERP responses to gain and loss feedback during this age (17). Although no other studies have investigated internal consistency of fMRI response to gain and loss feedback, it should be noted that the internal consistency of the ventral striatal response to gain reported here are higher than the poor to moderate test-retest reliability for ventral striatal responses to gain feedback reported in adult studies (13, 15). This pattern would be unsurprising for a measure with high internal consistency that is also sensitive to potential state-related changes across testing sessions.

Although other studies have demonstrated that feedback-related ERPs are internally consistent during late childhood and early adolescence (17), it was not clear whether internal consistency of neural response to gains and losses differ with age during this developmental stage. The current results suggest that ERP and BOLD responses to gain and loss feedback are similar in their true score from age 8 through 14 years. This is particularly important to establish, as studies investigating relationships between reward-related neural activity and individual differences are increasingly being conducted in pediatric groups (30, 43). The current results suggest that age-related differences in striatal response to reward during adolescence may not simply reflect variability in psychometric properties. Future studies, including participants of a wider age range, as well as males, are needed to replicate and extend the current findings.

Although the receipt of monetary gains versus losses elicits robust responses within the ventral striatum at the group level (44), it is possible that there is high trial-to-trial variability within subjects in the magnitude of those responses within subject, i.e. poor internal consistency. This is crucial insofar as internal consistency places an upper limit on validity, or the degree to which measures can correlate with one another. Poor internal consistency will weaken the ability of a measure to relate to other individual difference measures. In fact, this pattern was observed by Bress et al., 2015, where ERPs to gains had better psychometric properties and had a more *consistent* relationship with depressive symptoms across two testing sessions than difference-based measures of reward (17).

After establishing that response to gain and loss showed high internal consistency we investigated whether these responses also showed sufficient internal consistency when the first and second half of trials were analyzed separately. Both ERP and BOLD responses to gain and loss all had high internal consistency during both the first and second half of the task – indicating that a internally consistent measures of response to gain and loss in 8 to 14 year-old girls can be obtained using just 14 gain and 14 loss trials. This has important practical implications, as a task with fewer trials would require less time in the scanner. Although both ERP and VS responses showed similar internal consistency across the task, it is important to note that there was a marked decrease in the internal consistency of PFC responses, particularly for gains, from the first to the second half of the task. It is possible that decreases in PFC internal consistency reflect a decrease in general task engagement over

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

time. However, another interesting hypothesis is that decreasing internal consistency within the PFC reflects changes in the type of task engagement over time. Future work is needed to investigate this hypothesis and to further titrate the doors, and other reward tasks, and identify the fewest number of trials needed to provide neural measures of response to gain/loss with good psychometric properties while also maximizing relationships with individual differences. Further, as scanning parameters, e.g. TR length, may influence psychometric properties, future studies should also investigate internal consistency across a variety of scanning parameters.

In the current study, residual and subtraction based scores also had lower internal consistency across methodologies despite the relatively high internal consistency for responses to gain and loss. The lower internal consistency of difference scores has been discussed elsewhere and reflects accumulating measurement error (e.g., (45)). Internal consistency of subtraction-based scores is a particular concern when the constituent components show equal variance and are highly correlated (45, 46). For example, although responses to gain and loss were significantly correlated in this sample for all ROIs/ERPs and subtraction/residual scores were universally lower in their internal consistency, the lowest internal consistencies of the difference between gain and loss were observed for ROIs where the correlation between response to gain and loss was the highest (e.g. PFC). Residualized difference scores uniformly out-performed subtracted-based difference scores in terms of internal consistency in the current study, suggesting that this may be a more attractive option in individual differences research (39). Finally, we would note that despite *lower* internal consistency, difference scores can still isolate variance related to individual differences; this would be the case as long as the difference score captures the relevant individual differences variance.

Although the current study focused on internal consistency, another key issue is the construct validity of neural response to reward, the degree to which these measures relate to other constructs of interest. Although ERP and fMRI provide much different information on brain activity—in terms of both time-scale and mechanism, we previously found that the RewP correlated moderately with striatal response to reward in adults ($r = .52$ and $r = .28$ for left and right striatum, respectively (20)). In the current study, however, comparable correlations were much weaker (i.e., $r = .15$). One possibility is that large contextual differences in which ERP and fMRI data are acquired (i.e., supine body position; loud, anxiolytic environment in fMRI) may have a larger impact on the convergence of measures among adolescents relative to adults. That is, convergent validity between these measures may increase with age, if neural response to reward becomes less variable across contexts (i.e., trait-like) with age, this hypothesis should be investigated in the future. Future studies should also examine associations between both ERP and fMRI measures of reward in relation to development, self-report measures, and emerging symptoms of psychopatholgoy and risk—all of which will add to our understanding of construct validity of these neural measures.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

CONCLUSIONS

Both ERP- and fMRI-based measures of response to gain and loss during the doors task have high internal consistency that is invariant with respect to age; these measures are well suited for studying individual differences during emerging adolescence. Indeed, internal consistency was high for both ERP and fMRI measurements within the first half of the task. Both subtraction- and regression-based scores were characterized by lower internal consistency across ERP and fMRI methods. Regression-based residuals performed better, suggesting that this approach may be a preferable ‘difference-score’ method for individual differences research using this paradigm.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was supported by the National Institutes of Health [MH097767].

References

1. Insel T, Cuthbert B, Garvey M, Heinssen R, Pine DS, Quinn K, et al. Research domain criteria (RDoC): toward a new classification framework for research on mental disorders. The American journal of psychiatry. 2010; 167:748–751. [PubMed: 20595427]
2. Lilienfeld SO. The Research Domain Criteria (RDoC): an analysis of methodological and conceptual challenges. Behaviour research and therapy. 2014; 62:129–139. [PubMed: 25156396]
3. Cronbach LJ, Meehl PE. Construct validity in psychological tests. Psychological bulletin. 1955; 52:281–302. [PubMed: 13245896]
4. Meehl PE. Causes and effects of my disturbing little book. Journal of personality assessment. 1986; 50:370–375. [PubMed: 3806342]
5. Nelson BD, Perlman G, Klein DN, Kotov R, Hajcak G. Blunted Neural Response to Rewards Prospectively Predicts the Development of Depression in Adolescent Girls. American Journal of Psychiatry. 2016
6. Stice E, Yokum S, Burger KS. Elevated reward region responsivity predicts future substance use onset but not overweight/obesity onset. Biological psychiatry. 2013; 73:869–876. [PubMed: 23312561]
7. Silverman MH, Jedd K, Luciana M. Neural networks involved in adolescent reward processing: An activation likelihood estimation meta-analysis of functional neuroimaging studies. NeuroImage. 2015; 122:427–439. [PubMed: 26254587]
8. Galvan A, Hare TA, Parra CE, Penn J, Voss H, Glover G, et al. Earlier development of the accumbens relative to orbitofrontal cortex might underlie risk-taking behavior in adolescents. The Journal of neuroscience: the official journal of the Society for Neuroscience. 2006; 26:6885–6892. [PubMed: 16793895]
9. Luking KR, Luby JL, Barch DM. Kids, candy, brain and behavior: Age differences in responses to candy gains and losses. Developmental cognitive neuroscience. 2014; 9C:82–92.
10. Morgan JK, Olino TM, McMakin DL, Ryan ND, Forbes EE. Neural response to reward as a predictor of increases in depressive symptoms in adolescence. Neurobiology of disease. 2013; 52:66–74. [PubMed: 22521464]
11. Bress JN, Foti D, Kotov R, Klein DN, Hajcak G. Blunted neural response to rewards prospectively predicts depression in adolescent girls. Psychophysiology. 2013; 50:74–81. [PubMed: 23252717]
12. Fairchild G. The developmental psychopathology of motivation in adolescence. Developmental cognitive neuroscience. 2011; 1:414–429. [PubMed: 22436564]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

13. 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]
14. Wu CC, Samanez-Larkin GR, Katovich K, Knutson B. Affective traits link to reliable neural markers of incentive anticipation. NeuroImage. 2014; 84:279–289. [PubMed: 24001457]
15. Plichta MM, Schwarz AJ, Grimm O, Morgen K, Mier D, Haddad L, et al. Test-retest reliability of evoked BOLD signals from a cognitive-emotive fMRI test battery. NeuroImage. 2012; 60:1746–1758. [PubMed: 22330316]
16. Segalowitz SJ, Santesso DL, Murphy TI, Homan D, Chantziantoniou DK, Khan S. Retest reliability of medial frontal negativities during performance monitoring. Psychophysiology. 2010; 47:260–270. [PubMed: 20030755]
17. Bress JN, Meyer A, Proudfit GH. The stability of the feedback negativity and its relationship with depression during childhood and adolescence. Development and psychopathology. 2015; 27:1285–1294. [PubMed: 26439074]
18. Levinson A, Speed B, Hajcak G. (Under Review): Reliability of the electrocortical response to gains and losses in the Doors task. Psychophysiology.
19. Foti D, Weinberg A, Bernat EM, Proudfit GH. Anterior cingulate activity to monetary loss and basal ganglia activity to monetary gain uniquely contribute to the feedback negativity. Clinical neurophysiology: official journal of the International Federation of Clinical Neurophysiology. 2014
20. Carlson JM, Foti D, Mujica-Parodi LR, Harmon-Jones E, Hajcak G. Ventral striatal and medial prefrontal BOLD activation is correlated with reward-related electrocortical activity: a combined ERP and fMRI study. NeuroImage. 2011; 57:1608–1616. [PubMed: 21624476]
21. Bennett CM, Miller MB. How reliable are the results from functional magnetic resonance imaging? Annals of the New York Academy of Sciences. 2010; 1191:133–155. [PubMed: 20392279]
22. Vul E, Harris C, Winkielman P, Pashler H. Puzzlingly High Correlations in fMRI Studies of Emotion, Personality, and Social Cognition. Perspectives on psychological science: a journal of the Association for Psychological Science. 2009; 4:274–290. [PubMed: 26158964]
23. Lazar NA. Discussion of “Puzzlingly High Correlations in fMRI Studies of Emotion, Personality, and Social Cognition” by Vul et al. Perspectives on psychological science: a journal of the Association for Psychological Science. 2009; 4:308–309. 2009. [PubMed: 26158968]
24. Nichols TE, Poline JB. Commentary on Vul et al.’s (2009) “Puzzlingly High Correlations in fMRI Studies of Emotion, Personality, and Social Cognition”. Perspectives on psychological science: a journal of the Association for Psychological Science. 2009; 4:291–293. [PubMed: 26158965]
25. Braver TS, Cole MW, Yarkoni T. Vive les differences! Individual variation in neural mechanisms of executive control. Current opinion in neurobiology. 2010; 20:242–250. [PubMed: 20381337]
26. Nikolova YS, Bogdan R, Brigidi BD, Hariri AR. Ventral striatum reactivity to reward and recent life stress interact to predict positive affect. Biological psychiatry. 2012; 72:157–163. [PubMed: 22534456]
27. Plichta MM, Vasic N, Wolf RC, Lesch KP, Brummer D, Jacob C, et al. Neural hyporesponsiveness and hyperresponsiveness during immediate and delayed reward processing in adult attention-deficit/hyperactivity disorder. Biological psychiatry. 2009; 65:7–14. [PubMed: 18718573]
28. Forbes EE, Hariri AR, Martin SL, Silk JS, Moyles DL, Fisher PM, et al. Altered striatal activation predicting real-world positive affect in adolescent major depressive disorder. The American journal of psychiatry. 2009; 166:64–73. [PubMed: 19047324]
29. Forbes EE, Ryan ND, Phillips ML, Manuck SB, Worthman CM, Moyles DL, et al. Healthy adolescents’ neural response to reward: associations with puberty, positive affect, and depressive symptoms. Journal of the American Academy of Child and Adolescent Psychiatry. 2010; 49:162–172. e161–165. [PubMed: 20215938]
30. Luking K, Pagliaccio D, Luby J, Barch D. Depression Risk Predicts Blunted Responses To Candy Gains And Enhanced Responses To Candy Losses In Healthy Children. Journal of the American Academy of Child and Adolescent Psychiatry. 2016; 55:328–337. [PubMed: 27015724]
31. Braams BR, Peper JS, van der Heide D, Peters S, Crone EA. Nucleus accumbens response to rewards and testosterone levels are related to alcohol use in adolescents and young adults. Developmental cognitive neuroscience. 2016; 17:83–93. [PubMed: 26771250]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

32. Lancaster TM, Linden DE, Tansey KE, Banaschewski T, Bokde AL, Bromberg U, et al. Polygenic Risk of Psychosis and Ventral Striatal Activation During Reward Processing in Healthy Adolescents. JAMA Psychiatry. 2016; 73:852–861. [PubMed: 27384424]
33. Marco-Pallares J, Cucurell D, Munte TF, Strien N, Rodriguez-Fornells A. On the number of trials needed for a stable feedback-related negativity. Psychophysiology. 2011; 48:852–860. [PubMed: 21091959]
34. Speed BC, Nelson BD, Auerbach RP, Klein DN, Hajcak G. Depression risk and electrocortical reactivity during self-referential emotional processing in 8 to 14 year-old girls. Journal of abnormal psychology. 2016; 125:607–619. [PubMed: 27175985]
35. Weissman MM, Wickramaratne P, Gameroff MJ, Warner V, Pilowsky D, Kohad RG, et al. Offspring of Depressed Parents: 30 Years Later. The American journal of psychiatry. 2016 appiajp201615101327.
36. Proudfit GH. The reward positivity: from basic research on reward to a biomarker for depression. Psychophysiology. 2015; 52:449–459. [PubMed: 25327938]
37. Mazaika P, Hoeft F, Glover GH, Reiss AL. Methods and Software for fMRI Analysis for Clinical Subjects. Human brain mapping. 2009
38. Brett, M., Anton, J-L., Valabregue, R., Poline, J-B. International Conference on Functional Mapping of the Human Brain. Sendai, Japan: 2002. Region of interest analysis using an SPM toolbox.
39. Meyer A, Lerner M, Reyes A, Laird R, Hajcak G. Considering ERP difference scores as individual difference measures: Issues with subtraction and alternative approaches. Psychophysiology. in press.
40. Gratton G, Coles MG, Donchin E. A new method for off-line removal of ocular artifact. Electroencephalogr Clin Neurophysiol. 1983; 55:468–484. [PubMed: 6187540]
41. Hayes, AF. Introduction to Mediation, Moderation, and Conditional Process Analysis. New York, NY: Guilford Press; 2013.
42. Diedenhofen B, Musch J. Cocor: a comprehensive solution for the statistical comparison of correlations. PloS one. 2015; 10:e0121945. [PubMed: 25835001]
43. Luking K, Pagliaccio D, Luby J, Barch D. Reward Processing and Risk for Depression Across Development. Trends in cognitive sciences. 2016
44. Zhang WN, Chang SH, Guo LY, Zhang KL, Wang J. The neural correlates of reward-related processing in major depressive disorder: a meta-analysis of functional magnetic resonance imaging studies. Journal of affective disorders. 2013; 151:531–539. [PubMed: 23856280]
45. Chiou JS, Spreng RA. The reliability of difference scores: A re-examination. Journal of Consumer Satisfaction, Dissatisfaction, and Complaining Behavior. 1996; 9:158–167.
46. Gulliksen, H. Theory of Mental Tests. Hillsdale, NJ: Lawrence Erlbaum Associates Inc; 1987.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

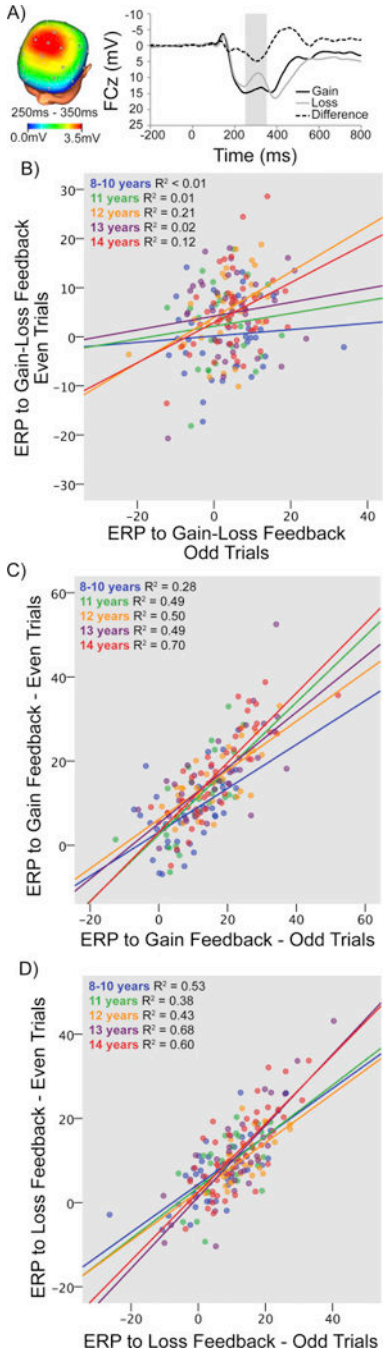

Figure 1. EEG Split-Half Relationships

A) Head map of the SCOREs (difference in ERP to gain-loss) and wave forms of the ERP to gain, loss, and difference at FCz. Internal consistency was weak for the B) SCOREs, but strong for the ERP to C) gain and D) loss. Split-half reliabilities did not differ with age.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

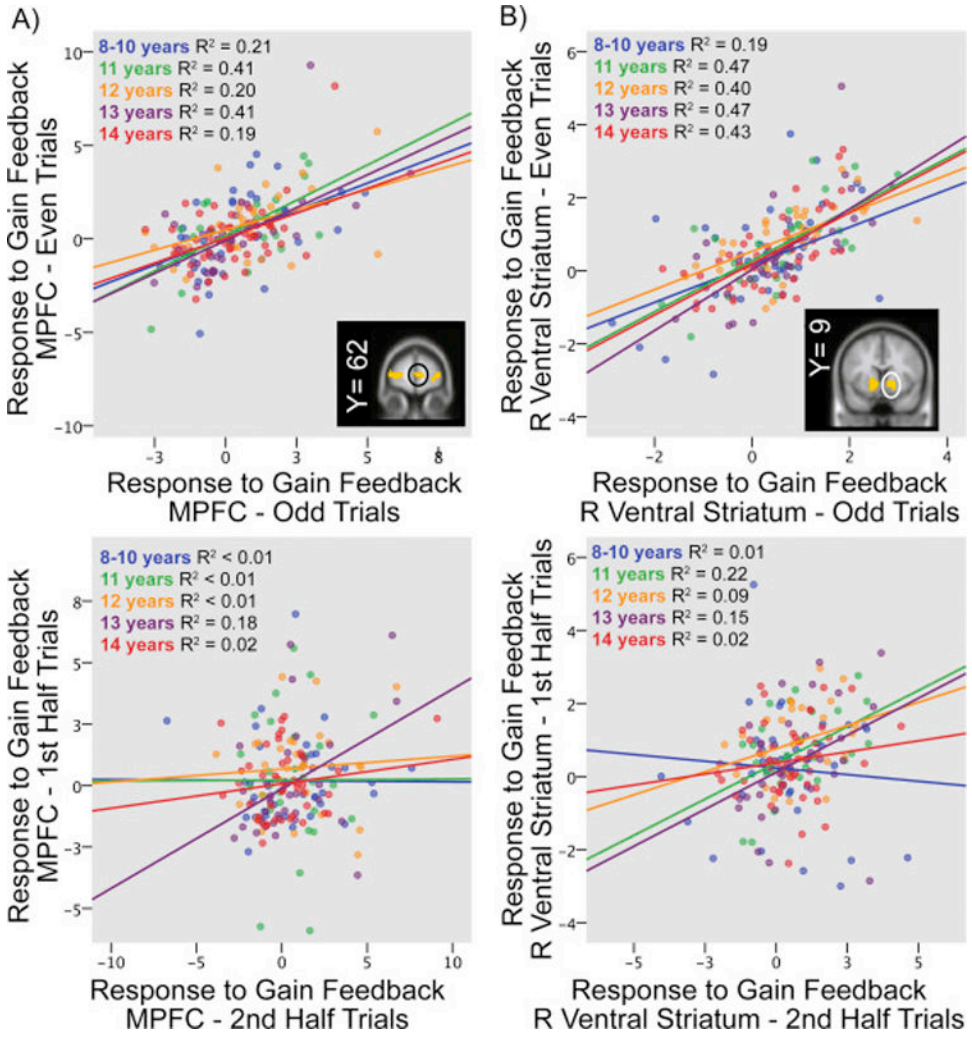

Figure 2. fMRI Split-Half Relationships by Age

For both the A) Medial Prefrontal Cortex (MPFC [6,63,3] black circle) and B) Right Ventral Striatum (R VS [13,7,−8] white circle) high internal consistency was observed for the response to gain feedback (top panels), but not for the difference in responses to gain versus loss feedback (gain-loss, bottom panels). Internal consistency did not differ with age for either region.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 1

Spearman Brown Coefficients for fMRI and EEG Split-Half Internal Consistency

Region	Gain	Loss	SCOREs	SCOREr
L Ventral Striatum	0.78	0.73	0.35	0.48
R Ventral Striatum	0.75	0.70	0.33	0.48
Medial PFC	0.69	0.66	0.02	0.27
L LPFC	0.76	0.72	0.20	0.36
R LPFC	0.80	0.75	0.04	0.26
ERP	0.85	0.86	0.36	0.50

Note: SCOREs = gain-loss, SCOREr = unstandardized residual gain controlling for loss, L=Left, R=Right, PFC=Prefrontal Cortex, LPFC=Lateral Prefrontal Cortex, ERP=Event Related Potential

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Table 2

Age Moderation of Split-Half Internal Consistency

Region	Gain × Age Interaction			Loss × Age Interaction			SCOREs × Age Interaction			SCOREr × Age Interaction		
	Coeff.	t	P	Coeff.	t	P	Coeff.	t	P	Coeff	t	P
L Ventral Striatum	0.08	2.16	0.03	<0.01	0.13	0.90	−0.02	−0.36	0.72	0.05	2.05	0.04
R Ventral Striatum	0.04	1.02	0.31	0.01	0.21	0.83	0.07	1.61	0.11	0.07	2.52	0.01
Medial PFC	−0.02	−0.47	0.64	0.02	0.71	0.48	−0.04	−1.13	0.26	−0.04	−0.88	0.38
L LPFC	−0.04	−1.19	0.23	<0.01	0.12	0.90	−0.05	−1.36	0.17	−0.03	−0.66	0.51
R LPFC	−0.01	−0.17	0.87	−0.01	−0.44	0.66	−0.05	−1.21	0.23	−0.01	−0.19	0.85
ERP	0.04	1.43	0.15	−0.04	−1.34	0.18	0.02	0.49	0.63	−0.03	−0.77	0.44

Note: SCOREs = gain-loss, SCOREr = unstandardized residual gain controlling for loss, L=Left, R=Right, PFC=Prefrontal Cortex, LPFC=Lateral Prefrontal Cortex, ERP=Event Related Potential. Correction for multiple comparisons, p<0.008.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Table 3

Split Half Relationships for the First and Second Half of Trials

	Gain			Loss			SCOREs			SCOREr		
	1st Half	2nd Half	Z (p)	1st Half	2nd Half	Z (p)	1st Half	2nd Half	Z (p)	1st Half	2nd Half	Z (p)
L V Striatum	0.77	0.67	1.95 (0.051)	0.71	0.64	1.20 (0.229)	0.37	0.44	−0.78 (0.434)	0.55	0.54	0.14 (0.894)
R V Striatum	0.72	0.67	0.90 (0.363)	0.64	0.66	−0.32 (0.747)	0.27	0.42	−1.59 (0.111)	0.50	0.52	−0.25 (0.801)
M PFC	0.71	0.53	2.77 (0.006)	0.71	0.59	1.95 (0.051)	0.44	0.13	3.19 (0.001)	0.55	0.37	2.14 (0.032)
L PFC	0.79	0.57	3.95 (<0.001)	0.70	0.64	1.02 (0.309)	0.32	0.27	0.51 (0.609)	0.55	0.44	1.36 (0.173)
R PFC	0.82	0.49	5.79 (<0.001)	0.74	0.66	1.47 (0.14)	0.38	0.13	2.51 (0.012)	0.58	0.24	3.90 (<0.001)
ERP	0.79	0.74	1.13 (0.259)	0.75	0.76	−0.22 (0.828)	0.21	0.35	−1.42 (0.156)	0.46	0.49	−0.36 (0.718)

Note: SCOREs = gain-loss, SCOREr = unstandardized residual gain controlling for loss, L=Left, R=Right, PFC=Prefrontal Cortex, LPFC=Lateral Prefrontal Cortex, ERP=Event Related Potential. Split half relationships within the first and second half of trials compared using corco.