International Journal of Psychophysiology 132 (2018) 323–330

Contents lists available at ScienceDirect

International Journal of Psychophysiology

journal homepage: www.elsevier.com/locate/ijpsycho

A longitudinal examination of event-related potentials sensitive to monetary reward and loss feedback from late childhood to middle adolescence

Autumn Kujawa[a,*], Ashley Carroll[a], Emma Mumper[b], Dahlia Mukherjee[a], Ellen M. Kessel[b], Thomas Olino[c], Greg Hajcak[d], Daniel N. Klein[b]

[a] Penn State College of Medicine, 22 Northeast Drive, Hershey, PA 17033, USA
[b] Stony Brook University, Stony Brook, NY 11794-2500, USA
[c] Temple University, 1701 North 13th Street, Philadelphia, PA 19122, USA
[d] Florida State University, 1107 West Call Street, Tallahassee, FL 32306-4301, USA

ARTICLE INFO

Keywords:
Electroencephalogram
Event-related potentials
Development
Reward
Reward positivity
Feedback negativity
Adolescence

ABSTRACT

Brain regions involved in reward processing undergo developmental changes from childhood to adolescence, and alterations in reward-related brain function are thought to contribute to the development of psychopathology. Event-related potentials (ERPs), such as the reward positivity (RewP) component, are valid measures of reward responsiveness that are easily assessed across development and provide insight into temporal dynamics of reward processing. Little work has systematically examined developmental changes in ERPs sensitive to reward. In this longitudinal study of 75 youth assessed 3 times across 6 years, we used principal components analyses (PCA) to differentiate ERPs sensitive to monetary reward and loss feedback in late childhood, early adolescence, and middle adolescence. We then tested reliability of, and developmental changes in, ERPs. A greater number of ERP components differentiated reward and loss feedback in late childhood compared to adolescence, but components in childhood accounted for only a small proportion of variance. A component consistent with RewP was the only one to consistently emerge at each of the 3 assessments. RewP demonstrated acceptable reliability, particularly from early to middle adolescence, though reliability estimates varied depending on scoring approach and developmental period. The magnitude of the RewP component did not significantly change across time. Results provide insight into developmental changes in the structure of ERPs sensitive to reward, and indicate that RewP is a consistently observed and relatively stable measure of reward responsiveness, particularly across adolescence.

1. Introduction

Processing of reward and loss feedback is essential to learning and shaping behaviors, and alterations in reward responsiveness likely play a role in the development of both internalizing and externalizing disorders (Zisner and Beauchaine, 2016). As such, there has been growing interest in the measurement of individual differences in reward responsiveness across levels of analysis (National Institute of Mental Health, 2017), including behavioral (Pizzagalli et al., 2005), circuit (Liu et al., 2011), and neurophysiological measures, such as event-related potentials (ERPs; Proudfit, 2015).

Brain circuits underlying reward processing undergo considerable development from childhood into adolescence, with evidence of differential patterns of maturation of subcortical regions, such as the striatum, and regions of the prefrontal cortex (PFC) involved in cognitive control. That is, compared to both children and adults, adolescents show heightened activation of the striatum during receipt of reward (Casey et al., 2008; Galvan, 2010; Shulman et al., 2016). On the other hand, top-down cognitive control regions, such as lateral PFC, are thought to continue to mature into adulthood and increase in activation from adolescence to adulthood (Casey et al., 2008; Galvan, 2010; Shulman et al., 2016).

To complement circuit measures of reward responsiveness, neurophysiological measures, such as ERPs, are economically and easily assessed across development and provide insight into the temporal dynamics of reward processing (Nelson and McCleery, 2008). In particular, an ERP component known as the reward positivity (RewP) or feedback negativity, is a relative positivity following receipt of a reward or positive feedback approximately 300 ms after feedback over frontocentral sites in youth and adults (Foti et al., 2011; Gehring and Willoughby, 2002). RewP appears to be a valid measure of individual differences in reward responsiveness. It has been shown to correlate

* Corresponding author at: Department of Psychiatry, Penn State College of Medicine, 22 Northeast Drive, Hershey, PA 17033, USA.
E-mail address: autumn.kujawa@gmail.com (A. Kujawa).

https://doi.org/10.1016/j.ijpsycho.2017.11.001
Received 28 June 2017; Received in revised form 1 November 2017; Accepted 3 November 2017
Available online 04 November 2017
0167-8760/

with activation in subcortical and cortical brain regions involved in reward processing, including ventral striatum, anterior cingulate cortex, and medial PFC (Becker et al., 2014; Carlson et al., 2011), as well as self-report and behavioral measures of reward sensitivity and positive emotionality (Bress and Hajcak, 2013; Kujawa et al., 2015). Moreover, altered reward responsiveness, as measured by RewP, appears to play a role in the emergence of psychopathology, particularly depression, in children and adolescents (e.g., Belden et al., 2016; Bress et al., 2013; Kujawa and Burkhouse, 2017; Nelson et al., 2016). Yet, the extent to which developmental changes in circuits underlying reward processing are reflected in the development of RewP or other ERP components sensitive to reward and loss feedback remains relatively unexplored.

In addition to validity, reliable measures of reward responsiveness are essential for examining developmental changes, correspondence across levels of analysis, and assocations with the emergence of psychiatric symptoms. In general, ERP amplitudes tend to be stable across time (Cassidy et al., 2012), with evidence that ERPs measured in children show comparable reliability to ERPs in adults (Hämmerer et al., 2013). Moreover, there is growing evidence that RewP is a reliable measure of reward responsiveness that shows good internal consistency and test-retest reliability (Bress et al., 2015; Levinson et al., 2017; Luking et al., 2017; Segalowitz et al., 2010). Specifically, in young adults assessed across one week, strong test-retest reliability was observed for RewP to losses and gains separately (*rs* = 0.45 and 0.71), with lower reliability for difference score measures (*rs* = 0.22 and 0.27; Levinson et al., 2017). One longitudinal study of 8- to 13-year-olds also found strong reliability for RewP to monetary losses and gains assessed across two years (*rs* = 0.64 and 0.67), but lower reliability of RewP as a difference score (*rs* = 0.18 to 0.29; Bress et al., 2015). Given the broad age range of this sample, the authors were unable to evaluate test-retest reliability across specific developmental periods (e.g., childhood into early adolescence), which may be particulary important for evaluating the utility of ERPs for examining the emergence of psychopathology.

Although there is evidence to indicate ERP measures of reward responsiveness demonstrate strong psychometric properties and are useful for informing understanding of the role of altered reward processing in the development of psychopathology, a number of gaps in the literature remain. First, within-subject, longitudinal work has yet to systematically evaluate typical developmental changes in ERP measures of reward responsiveness, including the timing and scalp distributions of these components at discrete developmental periods and both rank-order and mean-level stability. Second, although there is some evidence that RewP is reliably measured across development, it is unclear how reliability may be affected by specific developmental stages or whether test-retest reliability of RewP is maintained for developmental periods longer than 2 years. To further inform understanding of ERP measures of reward responsiveness and optimal methods across development, we first used principal components analyses (PCA; Dien, 2012) to systematically differentiate timing and spatial distributions of neural activity in response to monetary reward and loss feedback in a longitudinal sample of youth assessed at 3 time points, spanning a period of 6 years (i.e., late childhood, early adolescence, and middle adolescence). This approach enabled us to identify the underlying components of reward-related ERPs at each assessment and examine qualitative developmental changes. Next, for reward-related components emerging across development, we evaluated rank-order and mean-level stability and tested typical developmental changes in the magnitude of ERP responses to rewards and losses.

2. Materials and methods

2.1. Participants

Participants were part of a larger community sample of children initially recruited when the children were 3 or 6 years old (see Kujawa et al., 2014; Olino et al., 2010). Participants were invited back to the laboratory for electroencephalogram (EEG) assessments approximately every 3 years following the initial assessment. The current study included data from a subset of 75 participants who completed the monetary reward task at 3 time points between late childhood and middle adolescence. Data were available for 90 participants who completed the most recent assessment in middle adolescence. Of these, 5 participants were missing data from one of the previous assessments and 10 participants were excluded for excessive noise in the EEG data at 1 or more assessments, yielding the total sample of 75. Mean age of the sample was 9.40 (*SD* = 0.43) at the late childhood assessment, 13.05 (*SD* = 0.24) at the early adolescence assessment, and 15.16 (*SD* = 0.16) at the middle adolescence assessment. The sample was 44.0% female, 8.0% Hispanic/Latino, 97.3% Caucasian, 1.3% African American, and 1.3% Asian American. This study was approved by the Stony Brook University Institutional Review Board. Parents of participants provided informed consent and children provided assent.

2.2. Measures

2.2.1. Reward task

The EEG reward task has been used in previous studies to elicit the RewP (Bress and Hajcak, 2013; Bress et al., 2015; Kujawa et al., 2014). Participants were told they could win up to $5 and completed practice trials before beginning the task. The task consisted of 60 trials, presented in three blocks of 20 trials. At the beginning of each trial, participants were presented with an image of two doors and instructed to choose one door by clicking the left or right mouse button. The doors remained on the screen until the participant responded. Next, a fixation mark (+) appeared for 1000 ms, and feedback was presented on the screen for 2000 ms. Participants were told that they could either win $0.50 or lose $0.25 on each trial. A gain was indicated by a green "↑," and a loss was indicated by a red "↓." Finally, a fixation mark appeared again and was followed by the message "Click for the next round", which remained on the screen until the participant responded and the next trial began. Across the task, 30 gain and 30 loss trials were presented in a random order. Participants received $5 following completion of the task.

2.2.2. EEG data collection and processing

Continuous EEG was recorded at each assessment using a 34-electrode cap (32 channels with the addition of FCz and Iz) and a BioSemi system (BioSemi, Amsterdam, Netherlands). The electrooculogram (EOG) generated from eye movements and blinks was recorded using facial electrodes placed approximately 1 cm above and below the eye and 1 cm from the outer corners of the eyes. Electrodes were also placed on the left and right mastoids. Per the design of the BioSemi system, the common mode sense active electrode and driven right leg passive electrode served as the reference and ground electrodes during data acquisition. Recordings were digitized with a sampling rate of 1024 Hz.

Offline processing was conducted using BrainVision Analyzer software (Brain Products, Munich, Germany). Data were referenced to an average of the recordings from left and right mastoids, band-pass filtered with cutoffs of 0.01 and 30 Hz, and segmented for each trial 500 ms before feedback, continuing for 1000 ms after feedback onset. In cases of faulty recordings from a specific electrode, data were interpolated from surrounding electrodes. Eye-blink correction (Gratton et al., 1983) and semi-automatic artifact rejection procedures were conducted. Criteria of a voltage step of 50 μV between sample points, a maximum voltage difference of 300 μV within a 200 ms interval, and minimum activity of 0.5 μV within 100 ms intervals were used to automatically detect artifacts, with additional artifacts removed by visual inspection. All participants had a minimum of 15 segments per condition at Cz after artifact rejection, and the mean number of included segments per condition was 28.08 (*SD* = 2.61). ERPs were averaged for reward and loss feedback, and baseline corrected to activity 500 ms prior to feedback.

2.3. Data analysis

2.3.1. PCA

Temporospatial PCAs on averaged ERP data at each assessment were conducted using the ERP PCA Toolkit, version 2.54 (Dien, 2010b). First, temporal PCAs were conducted using all time points from each participant's averaged data as variables, and participants, trial types, and recording sites as observations. A Promax rotation was used to rotate to simple structure in the temporal domain (Dien, 2010a; Dien et al., 2007). A parallel analysis (Horn, 1965) was conducted on the resulting Scree plot (Cattell, 1966) in which the Scree of the actual dataset was compared to a Scree derived from a fully random dataset. The largest number of factors accounting for a greater proportion of variance than the fully random dataset were retained (Dien, 2010a). Based on this criterion, 10 temporal factors (TF) were retained at the late childhood assessment, 19 in early adolescence, and 14 in middle adolescence. Following the temporal PCA, a spatial PCA was conducted on each temporal factor (Dien, 2010a; Dien et al., 2007). Variables consisted of all recording sites, and observations included participants, trial types, and temporal factor scores. Infomax was used to rotate the spatial factors to independence (Dien, 2010a). Based on the results of the parallel test, 3 spatial factors (SF) were extracted from each TF in late childhood, and 2 SF were extracted from each TF at each of the adolescent assessments. Temporospatial PCAs resulted in 30 factor combinations at the initial assessment, 38 at the early adolescence assessment, and 28 at the middle adolescence assessment that combined accounted for 33.0 to 35.8% unique variance in the data at each assessment.

In order to focus further analyses on components that significantly differentiate reward and loss feedback and account for a meaningful proportion of variance in the ERP wave, robust analysis of variance (ANOVA; Dien, 2017; Keselman et al., 2003) with 4999 bootstrapping simulations was conducted in ERP PCA Toolkit on PCA factors converted to microvolt scaling (Dien, 2012) for components that accounted for a minimum of 0.5% of unique variance. Given potential variability in p values using this approach, simulations were run 11 times, with median p values reported, and only results in which the median p value plus 2 standard deviations remained below 0.05 were considered significant (Dien, 2017). Lastly, Bonferroni correction was applied to analyses at each assessment to correct for multiple comparisons.

2.3.2. Reliability and developmental change analyses

Stability and change over time was evaluated for ERP components emerging across all 3 assessments. Two approaches to scoring were used to evaluate reliability and developmental changes: microvolt-scaled factor scores derived from ERP PCA Toolkit and mean activity in microvolts extracted from ERP averages using BrainVision Analyzer. Mean activity was evaluated to be consistent with more traditional and commonly used scoring approaches in ERP research (Luck, 2005) and to evaluate a scoring method that is not dependent on the PCA solution. Analyses of bivariate (Pearson's r) and intraclass correlations (ICC) were computed in SPSS 23. Pearson's r assesses rank-order stability between two assessments, whereas ICC provides a measure of both rank-order and mean-level consistency between two assessments and across all three assessments. Two-way mixed single measures ICCs with absolute agreement were calculated. We examined reliability of the ERP response to reward and loss separately, as well as two common difference score approaches for isolating the variance of interest: subtraction-based and residual scores (Luck, 2005; Meyer et al., 2017). Finally, robust ANOVAs on mean activity were computed in ERP PCA Toolkit to evaluate developmental changes in the magnitude of ERP responses to monetary reward and loss feedback.

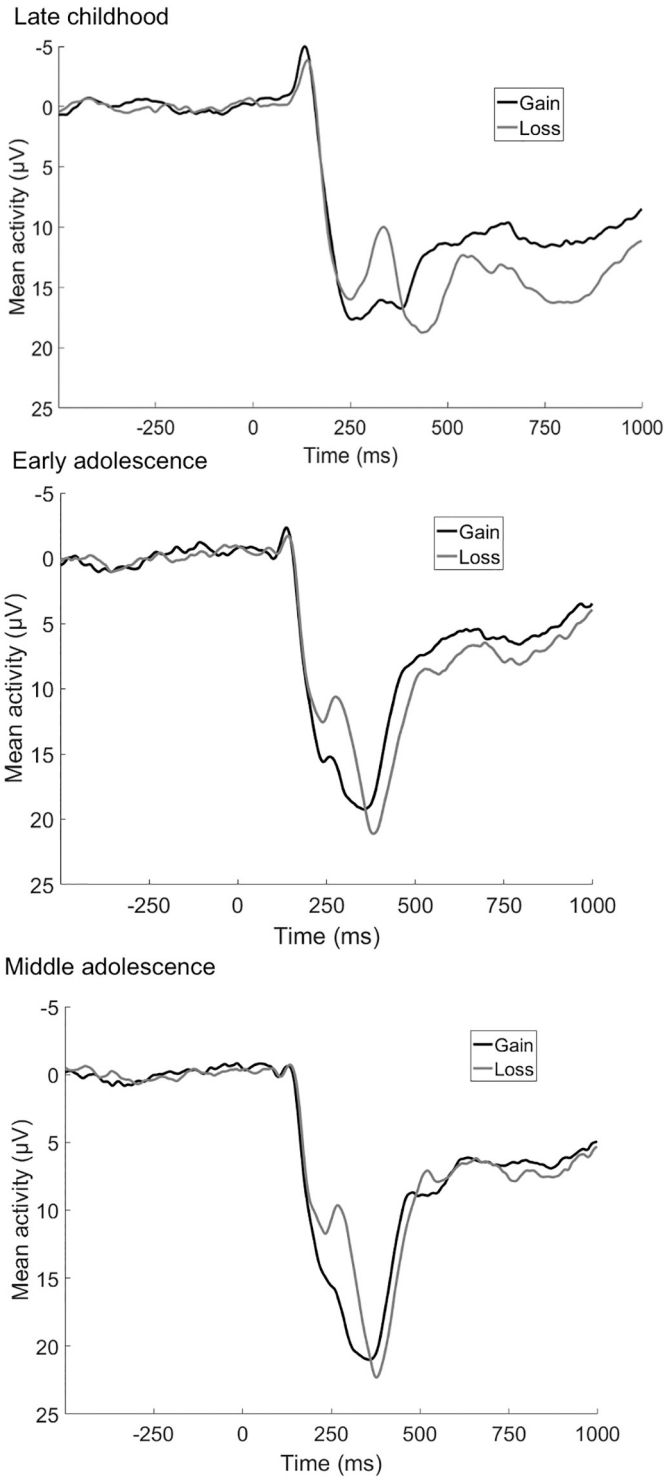

Fig. 1. Average ERP responses to gains and losses at Cz prior to PCA.

3. Results

3.1. PCA of ERPs sensitive to reward and loss at each assessment

ERP waves at Cz for each assessment prior to PCA are presented in Fig. 1. Factor combinations that accounted for a minimum of 0.5% unique variance in the ERP wave at each assessment and significantly differentiated rewards from losses are described in Table 1 (descriptions refer to whether the component was relatively more positive or negative in response to reward or loss feedback).

Table 1
Temporospatial factor combinations sensitive to monetary gain vs. loss feedback at each assessment (descriptions indicate whether the component was relatively more positive or negative in response to gain vs. loss feedback).

Late childhood					
Factor combination	Unique variance (%)	Temporal peak (ms)	Peak electrode	Description	Gain vs. loss (*T*)
TF8/SF1	0.54	106	FCz	Frontocentral positivity for loss	14.93**
TF7/SF1	0.90	147	Cz	Central negativity for loss	15.98**
TF5/SF1	0.77	284	Cz	Central positivity for gain	31.31**
TF10/SF1	0.57	497	FCz	Frontocentral negativity for gain	16.47**
TF6/SF1	0.71	999	FP1	Frontoparietal positivity for gain	11.76*
Early adolescence					
Factor combination	Unique variance (%)	Temporal peak (ms)	Peak electrode	Description	Gain vs. loss (*T*)
TF2/SF1	5.05	308	Cz	Central positivity for gain	27.60**
TF3/SF1	2.81	461	Cz	Central positivity for loss	12.92*
Middle adolescence					
Factor combination	Unique variance (%)	Temporal peak (ms)	Peak electrode	Description	Gain vs. loss (*T*)
TF9/SF1	0.69	197	Cz	Central negativity for gain	25.67**
TF2/SF1	6.65	310	Cz	Central positivity for gain	49.54**
TF4/SF1	1.74	387	FCz	Central positivity for loss	48.60**

T values from robust ANOVA; TF = temporal factor; SF = spatial factor; *$p < 0.01$; **$p < 0.001$.

At each assessment, a relative positivity was identified that was enhanced for reward compared to loss feedback (i.e., RewP), peaking between 284 and 310 ms after feedback onset at Cz (TF5/SF1 in late childhood; TF2/SF1 in early adolescence and middle adolescence; Table 1). ERP waves and scalp distributions for components corresponding to RewP derived from PCA at each assessment are presented in Fig. 2.

At each of the adolescent assessments, a later positivity was observed that was enhanced for loss compared to reward feedback and peaked between 387 and 461 ms after feedback at Cz in early adolescence and the adjacent FCz in middle adolescence, possibly reflecting the P300 component. A similar component was not observed to significantly differentiate responses to reward and loss feedback in late childhood. Four additional positive and negative ERP components significantly differentiated responses to reward vs. loss feedback in late childhood; however, each component accounted for a small proportion of variance in the ERP. As only the component consistent with RewP emerged across each of the assessments, statistical analyses of reliability and developmental changes focused on this component.

3.2. Test-retest reliability of RewP

Reliability and developmental changes in RewP were evaluated using 2 ERP scoring approaches: PCA factor scores converted to microvolt scaling for the component corresponding to RewP at each assessment and mean activity in microvolts 250–350 ms at FCz/Cz, which is consistent both with the peaks identified by PCA and where RewP tends to be most apparent (e.g., Nelson et al., 2016; Proudfit, 2015).[1]

3.2.1. PCA factor score measures

Descriptive statistics, bivariate correlations, and ICCs for PCA factor score measures of RewP are presented in Table 2. RewP to gains and RewP to losses showed moderate to strong rank-order stability (Pearson's *r*) across assessments, though correlations were more modest for difference score measures (subtraction-based and residuals). With regard to mean-level/rank-order stability, ICCs across all 3 assessments did not reach the level considered acceptable for psychological assessments (≥ 0.40; Cicchetti, 1994); however, ICCs for all RewP PCA factor scores, including difference scores, indicated fair to good reliability from early to middle adolescence.

3.2.2. Mean activity measures

Descriptive statistics, bivariate correlations, and ICCs for mean activity measures of RewP are presented in Table 3. RewP to gains and RewP to losses showed strong rank-order stability (Pearson's *r*) across assessments. Pearson's correlations were more modest for difference score measures (subtraction-based and residuals), but remained moderate in magnitude for most associations. With regard to mean-level/rank-order stability, ICCs across all 3 assessments indicated fair reliability for RewP to gains and losses, but ICCs for difference score measures fell below the level considered acceptable for psychological assessments (Cicchetti, 1994). ICCs for difference scores were relatively higher for early to middle adolescence, with fair reliability of residual scores observed in this developmental period.

3.3. Developmental change in RewP magnitude

Lastly, 3 (assessment: late childhood, early adolescence, middle adolescence) × 2 (feedback: gain vs. loss) robust ANOVAs were computed in ERP PCA Toolkit to evaluate developmental changes in the magnitude of RewP to gain and loss feedback. As can be observed in Fig. 2, with RewP evaluated using PCA factor scores, the effect of age on microvolt-scaled PCA factor scores for the component corresponding to RewP at each assessment was significant, $T_{WJt}/c(2, 60) = 90.75$, $p < 0.001$, such that the scores increased from late childhood to adolescence, but this effect was not moderated by feedback type ($p = 0.38$). Moreover, this effect was not reflected in ERP averages before PCA. That is, with RewP scored as mean activity 250–350 ms at FCz/Cz, the main effect of age and age X feedback interaction were not significant (*ps* > 0.43).

4. Discussion

In this longitudinal study, we used PCA to differentiate the timing and spatial distributions of ERPs sensitive to monetary reward feedback, and examined reliability of and developmental changes in ERP measures of reward responsiveness from late childhood to middle

[1] RewP was also evaluated as mean activity at Cz in the 100 ms window around the PCA-defined peak at each assessment and comparable results were obtained, indicating that reliability is similar when adjusting the scoring of RewP based on the specific peaks identified by PCA (see note for Early Adolescence).

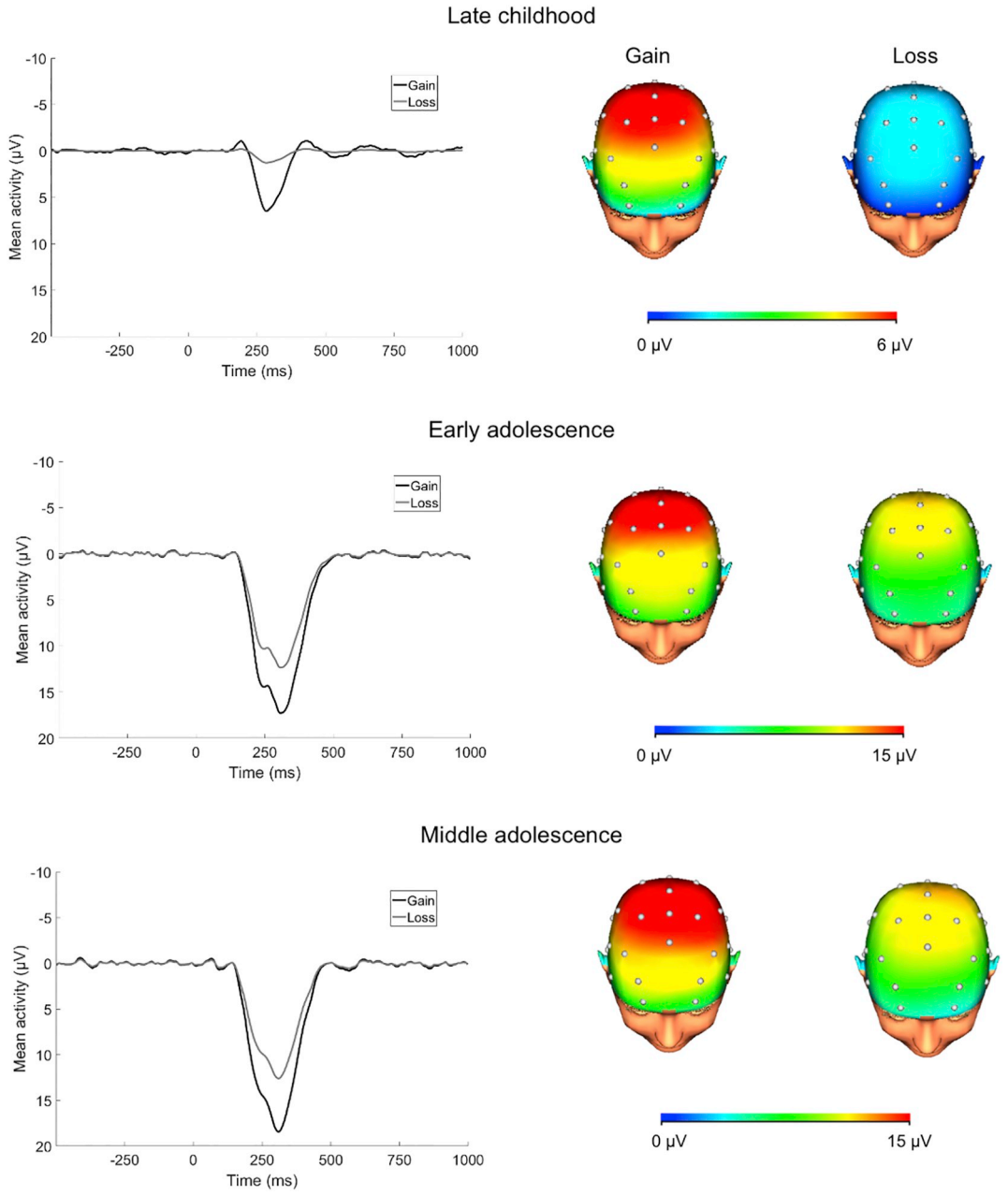

Fig. 2. ERPs and scalp distributions of the PCA temporospatial factor combinations corresponding to RewP at each assessment. Note: scalp distributions for the late childhood assessment are presented on a different scale than the other two assessments.

Table 2
Means, standard deviations, bivariate correlations (Pearson's *r*), and intraclass correlations (ICC) of the PCA factor scores for the component corresponding to the reward positivity (RewP) at each assessment.

	T1 *M*(SD)	T2 *M*(SD)	T3 *M*(SD)	T1 to T2 *r*	T2 to T3 *r*	T1 to T3 *r*	T1 to T2 ICC	T2 to T3 ICC	T1 to T3 ICC	T1, T2, and T3 ICC
PCA factor gain	6.48(7.42)	17.27(10.00)	18.44(10.88)	0.46***	0.66***	0.29*	0.25***	0.65***	0.15**	0.32***
PCA factor loss	1.26(6.39)	12.33(9.09)	12.61(8.37)	0.34**	0.58***	0.31**	0.16**	0.59***	0.14**	0.25***
PCA factor difference	5.21(7.87)	4.94(7.55)	5.83(7.60)	0.26*	0.43***	0.27*	0.26*	0.43***	0.27*	0.32***
PCA factor residuals	0.00(6.93)	0.00(7.23)	0.00(7.58)	0.26*	0.48***	0.15	0.26*	0.48***	0.15	0.30***

T1 = late childhood; T2 = early adolescence; T3 = middle adolescence; PCA = principal components analysis; ***$p < 0.001$; **$p < 0.01$; *$p < 0.05$.

adolescence. PCA results identified qualitative differences in ERP components sensitive to reward in late childhood relative to adolescence; however, a component consistent with the RewP (i.e., a relative positivity in response to rewards compared to losses maximal around 300 ms at central recording sites) consistently emerged at each of the 3 assessments. Moreover, quantitative comparisons of the magnitude of RewP at each assessment suggested that this component is relatively stable across development, particularly from early to middle adolescence.

PCA results indicated that reward and loss feedback modulate a

Table 3
Means, standard deviations, bivariate correlations (Pearson's *r*), and intraclass correlations (ICC) of the reward positivity (RewP) scored as mean activity 250–350 ms after feedback at FCz/Cz at each assessment.

	T1 *M*(SD)	T2 *M*(SD)	T3 *M*(SD)	T1 to T2 *r*	T2 to T3 *r*	T1 to T3 *r*	T1 to T2 ICC	T2 to T3 ICC	T1 to T3 ICC	T1, T2, and T3 ICC
RewP gain	14.73(9.15)	15.71(9.48)	16.96(10.47)	0.62***	0.62***	0.52***	0.62***	0.61***	0.51***	0.58***
RewP loss	10.60(8.01)	10.62(8.87)	11.31(7.76)	0.53***	0.57***	0.53***	0.53***	0.57***	0.53***	0.54***
RewP difference	4.14(7.24)	5.09(6.75)	5.65(7.01)	0.21	0.39**	0.36**	0.21*	0.39***	0.35**	0.32***
RewP residuals	0.00(6.95)	0.00(6.47)	0.00(7.01)	0.30**	0.43***	0.34**	0.31**	0.43***	0.35**	0.36***

T1 = late childhood; T2 = early adolescence; T3 = middle adolescence; RewP = reward positivity; ***$p < 0.001$; **$p < 0.01$; *$p < 0.05$.

greater number of ERP components in childhood than in adolescence, though in childhood, these components accounted for a small proportion of variance. There are a few possible explanations for these results. First, imaging data indicate that brain regions activated during cognitive tasks become less diffuse from childhood to adolescence, such that children appear to recruit regions that are less essential to task performance, with patterns of activation becoming more focal into adolescence (Casey et al., 2008; Durston et al., 2006). As such, the current ERP findings suggest the possibility that the temporal dynamics of feedback processing may also be more diffuse in childhood, becoming more integrated into adolescence. It is also important to note that EEG data collected from children tends to have lower signal-to-noise ratios than data collected from adolescents and adults (Hämmerer et al., 2013). Although the same artifact rejection and data cleaning procedures were used at each assessment, it is possible that additional noise in the data at the childhood assessment partly contributed to the emergence of a greater number of components accounting for small proportions of variance. The current findings highlight the importance of multiple assessments in developmental research, along with the use of approaches like PCA in order to systematically identify components that consistently emerge across development and avoid drawing conclusions based on ERP modulations that could be driven by artifacts (Dien, 2012).

Though distinct patterns of ERPs emerged in the PCA at each assessment, a component corresponding to RewP consistently emerged at each developmental stage with very similar temporal and spatial features. Moreover, this component exhibited acceptable rank-order and mean-level reliability when scored as mean activity in response to gains and losses, even over 6 years and across a developmental period marked by significant neurodevelopmental changes (Casey et al., 2008; Shulman et al., 2016). When examined across all 3 assessments, reliability of PCA and difference score measures was lower than that observed for mean activity in response to gains and loss separately. This reduction in reliability appeared to be driven at least in part by developmental differences in late childhood, as fair to good reliability was observed for both PCA and difference score measures across early and middle adolescence.

PCA factor scores provide a measure of RewP removing any potential overlap with other components. As such, it would be expected that ERP measures derived from PCA may show better reliability than traditional scoring approaches. However, these scores are dependent on the specific PCA solution at each assessment, which changed across development. Our results indicate that PCA factor scores may be more reliable than traditional scoring approaches when the PCA factor solution is similar across assessments, but when the structure of ERP components changes across time, mean activity measures are likely to be more reliable. That is, PCA factors exhibited fair to good reliability across early and middle adolescence but not from late childhood to adolescence, consistent both with developmental changes in reward-related brain function in adolescence (Casey et al., 2008; Shulman et al., 2016) and qualitative differences observed using PCA in the current study.

The relatively low reliability of difference scores is not surprising, as similar patterns have been previously observed for internal consistency and rank-order stability of RewP (Bress et al., 2015; Levinson et al., 2017; Luking et al., 2017). Across areas of research, reliability of difference scores is affected by noise and error in both measures and is typically lower than the average reliability of the two individual measures (Furr and Bacharach, 2013). Moreover, ERP components tend to be highly correlated with each other (e.g., Levinson et al., 2017), which has been shown to produce difference scores that are lower in reliability (Furr and Bacharach, 2013). Difference scores are often preferred in ERP research in order to isolate the variance in the ERP wave attributed to the cognitive process of interest (Luck, 2005), and there is some evidence that difference score measures of RewP may relate more strongly to certain individual differences than ERPs in a single condition (e.g., Kujawa et al., 2014). Similar to PCA factor scores, reliability of difference scores depended on developmental timing, with acceptable reliability from early to middle adolescence but relatively lower reliability when including the late childhood assessment. Lastly, similar to previous observations (Luking et al., 2017), our analyses indicate that residual approaches to difference scores may be somewhat more reliable than subtraction-based approaches.

Surprisingly, RewP was the only component that differentiated reward and loss feedback and consistently emerged at each assessment. Though a component similar to P300 and enhanced for losses compared to wins emerged at both of the adolescent assessments, this component was not observable at the childhood assessment, suggesting that it may emerge later in development. In addition, an N200 component has been observed in adults in response to losses and is thought to either be suppressed by positive feedback or to overlap with the RewP to gains (Holroyd et al., 2008; Proudfit, 2015). However, similar to a previous PCA study in adults (Foti et al., 2011), we did not find evidence of a distinct N200 component modulated by reward and loss feedback in children and adolescents. Taken together, these results indicate that RewP is consistently and reliably measured from late childhood through middle adolescence. Although other ERP components appear to be modulated by reward and loss feedback in youth, the timing and scalp distributions of these components seems to vary, especially from late childhood to adolescence, with more consistency observed from early to middle adolescence.

Lastly, it is surprising that we did not find evidence that the magnitude of the RewP changed significantly across development in this sample, particularly given evidence of a developmental increase in activation of the striatum to rewards in adolescence (Shulman et al., 2016). Although several studies have linked RewP in adults to activation in the ventral striatum, ERPs are generated by large groups of neurons with limited spatial resolution, and RewP likely reflects activation of a broader network that includes regions of the PFC and anterior cingulate cortex (Becker et al., 2014; Carlson et al., 2011). Thus, RewP may be less sensitive to developmental changes than more direct measures of striatal activation. However, longitudinal examination of developmental changes in the magnitude of RewP warrants study in larger samples continuing into later adolescence. Differences in the PCA solutions and reliability of RewP depending on the developmental period suggest that there are likely important developmental changes happening during this time, even though the overall magnitude of RewP did not significantly differ across assessments in this sample.

These results highlight a number of challenges to the field and further directions for future research. First, we relied on guidelines from the psychological assessment literature to interpret ICCs, but further research is needed to identify reliability standards for neural and physiological measures. Second, additional work is needed to evaluate standards for reliability, particularly mean-level stability, in longitudinal studies that span periods when developmental changes in brain function are expected. Relatedly, future work should continue to optimize ERP scoring methods across development. That is, PCA has clear benefits for identifying ERP components and removing effects of overlapping components, which is particularly critical given limited data on how ERPs change across development, but is limited by developmental changes in the structure of ERP components. As such, for developmental studies in which the PCA solution changes across time, a combined approach applying PCA to characterize components of interest, then incorporating traditional mean activity scoring for examination of individual differences may be most useful.

5. Conclusions

This longitudinal study is the first to evaluate ERP measures of reward responsiveness across 6 years of development and 3 assessments. Taken together, results indicate that although the structure of ERP components sensitive to monetary reward and loss change from childhood to middle adolescence, RewP is consistently observed and appears to be a relatively stable neurophysiological measure of reward responsiveness, particularly across adolescence. Results suggest that RewP is likely to be a useful ERP component for examining alterations in monetary reward responsiveness in the development of psychopathology.

Funding source

This work was supported by National Institute of Mental Health Grant R01 MH069942 to DNK.

Conflicts of interest

None.

References

Becker, M.P.I., Nitsch, A.M., Miltner, W.H.R., Straube, T., 2014. A single-trial estimation of the feedback-related negativity and its relation to BOLD responses in a time-estimation task. J. Neurosci. 34 (8), 3005–3012. http://dx.doi.org/10.1523/JNEUROSCI.3684-13.2014.

Belden, A.C., Irvin, K., Hajcak, G., Kappenman, E.S., Kelly, D., Karlow, S., ... Barch, D.M., 2016. Neural correlates of reward processing in depressed and healthy preschool-age children. J. Am. Acad. Child Adolesc. Psychiatry 55 (12), 1081–1089. http://dx.doi.org/10.1016/j.jaac.2016.09.503.

Bress, J.N., Hajcak, G., 2013. Self-report and behavioral measures of reward sensitivity predict the feedback negativity. Psychophysiology 50 (7), 610–616. http://dx.doi.org/10.1111/psyp.12053.

Bress, J.N., Foti, D., Kotov, R., Klein, D.N., Hajcak, G., 2013. Blunted neural response to rewards prospectively predicts depression in adolescent girls. Psychophysiology 50 (1), 74–81. http://dx.doi.org/10.1111/j.1469-8986.2012.01485.x.

Bress, J.N., Meyer, A., Proudfit, G.H., 2015. The stability of the feedback negativity and its relationship with depression during childhood and adolescence. Dev. Psychopathol. 17, 1285–1294.

Carlson, J.M., Foti, D., Mujica-Parodi, L.R., Harmon-Jones, E., Hajcak, G., 2011. Ventral striatal and medial prefrontal BOLD activation is correlated with reward-related electrocortical activity: a combined ERP and fMRI study. NeuroImage 57 (4), 1608–1616. http://dx.doi.org/10.1016/j.neuroimage.2011.05.037.

Casey, B.J., Jones, R.M., Hare, T.A., 2008. The adolescent brain. Ann. N. Y. Acad. Sci. 1124, 111–126. http://dx.doi.org/10.1196/annals.1440.010.

Cassidy, S.M., Robertson, I.H., O'Connell, R.G., 2012. Retest reliability of event-related potentials: evidence from a variety of paradigms. Psychophysiology 49 (5), 659–664. http://dx.doi.org/10.1111/j.1469-8986.2011.01349.x.

Cattell, R.B., 1966. The screen test for the number of factors. Multivar. Behav. Res. 1, 245–276.

Cicchetti, D.V., 1994. Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. Psychol. Assess. 6 (4), 284–290. http://dx.doi.org/10.1037/1040-3590.6.4.284.

Dien, J., 2010a. Evaluating two-step PCA of ERP data with Geomin, Infomax, Oblimin, Promax, and Varimax rotations. Psychophysiology 47 (1), 170–183. http://dx.doi.org/10.1111/j.1469-8986.2009.00885.x.

Dien, J., 2010b. The ERP PCA toolkit: an open source program for advanced statistical analysis of event-related potential data. J. Neurosci. Methods 187 (1), 138–145. http://dx.doi.org/10.1016/j.jneumeth.2009.12.009.

Dien, J., 2012. Applying principal components analysis to event-related potentials: a tutorial. Dev. Neuropsychol. 37 (6), 497–517. http://dx.doi.org/10.1080/87565641.2012.697503.

Dien, J., 2017. Best practices for repeated measures ANOVAs of ERP data: reference, regional channels, and robust ANOVAs. Int. J. Psychophysiol. 111, 42–56. http://dx.doi.org/10.1016/j.ijpsycho.2016.09.006.

Dien, J., Khoe, W., Mangun, G.R., 2007. Evaluation of PCA and ICA of simulated ERPs: Promax vs. infomax rotations. Hum. Brain Mapp. 28 (8), 742–763. http://dx.doi.org/10.1002/hbm.20304.

Durston, S., Davidson, M.C., Tottenham, N., Galvan, A., Spicer, J., Fossella, J.A., Casey, B.J., 2006. A shift from diffuse to focal cortical activity with development. Dev. Sci. 9 (1), 1–8. http://dx.doi.org/10.1111/j.1467-7687.2005.00454.x.

Foti, D., Weinberg, A., Dien, J., Hajcak, G., 2011. Event-related potential activity in the basal ganglia differentiates rewards from nonrewards: Temporospatial principal components analysis and source localization of the feedback negativity. Hum. Brain Mapp. 32 (12), 2207–2216. http://dx.doi.org/10.1002/hbm.21182.

Furr, R.M., Bacharach, V.R., 2013. Psychometrics: an Introduction, 2nd ed. Sage.

Galvan, A., 2010. Adolescent development of the reward system. Front. Hum. Neurosci. 4, 1–9. http://dx.doi.org/10.3389/neuro.09.006.2010.

Gratton, G., Coles, M.G.M.G., Donchin, E., 1983. A new method for off-line removal of ocular artifact. Electroencephalogr. Clin. Neurophysiol. 55 (4), 468–484. http://dx.doi.org/10.1016/0013-4694(83)90135-9.

Gehring, W.J., Willoughby, A.R., 2002. The medial frontal cortex and the rapid processing of monetary gains and losses. Science 295 (5563), 2279–2282. http://dx.doi.org/10.1126/science.1066893.

Hämmerer, D., Li, S., Völkle, M., Müller, V., 2013. A lifespan comparison of the reliability, test-retest stability, and signal-to-noise ratio of event-related potentials assessed during performance monitoring. Psychophysiology 50 (1), 111–123.

Holroyd, C.B., Pakzad-Vaezi, K.L., Krigolson, O.E., 2008. The feedback correct-related positivity: sensitivity of the event-related brain potential to unexpected positive feedback. Psychophysiology 45 (5), 688–697. http://dx.doi.org/10.1111/j.1469-8986.2008.00668.x.

Horn, J.L., 1965. A rationale and test for the number of factors in factor analysis. Psychometrika 30 (2), 179–185. http://dx.doi.org/10.1007/bf02289447.

Keselman, H.J., Wilcox, R.R., Lix, L.M., 2003. A generally robust approach to hypothesis testing in independent and correlated groups designs. Psychophysiology 40 (4), 586–596. http://dx.doi.org/10.1111/1469-8986.00060.

Kujawa, A., Burkhouse, K.L., 2017. Vulnerability to depression in youth: insights from affective neuroscience. Biol. Psychiatry 2 (1), 28–37.

Kujawa, A., Proudfit, G.H., Klein, D.N., 2014. Neural reactivity to rewards and losses in offspring of mothers and fathers with histories of depressive and anxiety disorders. J. Abnorm. Psychol. 123 (2), 287–297. http://dx.doi.org/10.1037/a0036285.

Kujawa, A., Proudfit, G.H., Kessel, E.M., Dyson, M., Olino, T., Klein, D.N., 2015. Neural reactivity to monetary rewards and losses in childhood: longitudinal and concurrent associations with observed and self-reported positive emotionality. Biol. Psychol. 104, 41–47. http://dx.doi.org/10.1016/j.biopsycho.2014.11.008.

Levinson, A.R., Speed, B.C., Infantolino, Z.P., Hajcak, G., 2017. Reliability of the electrocortical response to gains and losses in the doors task. Psychophysiology 54 (4), 601–607. http://dx.doi.org/10.1111/psyp.12813.

Liu, X., Hairston, J., Schrier, M., Fan, J., 2011. Common and distinct networks underlying reward valence and processing stages: a meta-analysis of functional neuroimaging studies. Neurosci. Biobehav. Rev. 35 (5), 1219–1236. http://dx.doi.org/10.1016/j.neubiorev.2010.12.012.

Luck, S.J., 2005. An Introduction to the Event-Related Potential Technique. The MIT Press, Cambridge, MA.

Luking, K.R., Nelson, B.D., Infantolino, Z.P., Sauder, C.L., Hajcak, G., 2017. Internal consistency of functional magnetic resonance imaging and electroencephalography measures of reward in late childhood and early adolescence. Biol. Psychiatry 2 (3), 289–297. http://dx.doi.org/10.1016/j.bpsc.2016.12.004.

Meyer, A., Lerner, M.D., de los Reyes, A., Laird, R.D., Hajcak, G., 2017. Considering ERP difference scores as individual difference measures: issues with subtraction and alternative approaches. Psychophysiology 54, 114–122. http://dx.doi.org/10.1111/psyp.12664.

National Institute of Mental Health, 2017. RDoC Matrix. Retrieved July 10, 2017, from. http://www.nimh.nih.gov/research-priorities/rdoc/constructs/rdoc-matrix.shtml.

Nelson, C.A., McCleery, J.P., 2008. Use of event-related potentials in the study of typical and atypical development. J. Am. Acad. Child Adolesc. Psychiatry 47 (11), 1252–1261. http://dx.doi.org/10.1097/CHI.0b013e18185a6d8.

Nelson, B.D., Perlman, G., Klein, D.N., Kotov, R., Hajcak, G., 2016. Blunted neural response to rewards prospectively predicts the development of depression in adolescent girls. Am. J. Psychiatry 173 (12), 1223–1230.

Olino, T.M., Klein, D.N., Dyson, M.W., Rose, S.A., Durbin, C.E., 2010. Temperamental emotionality in preschool-aged children and depressive disorders in parents: associations in a large community sample. J. Abnorm. Psychol. 119 (3), 468–478. http://

dx.doi.org/10.1037/a0020112.
Pizzagalli, D.A., Jahn, A.L., Shea, J.P.O., O'Shea, J.P., 2005. Toward an objective characterization of an anhedonic phenotype: a signal-detection approach. Biol. Psychiatry 57 (4), 319–327. http://dx.doi.org/10.1016/j.biopsych.2004.11.026.
Proudfit, G.H., 2015. The reward positivity: from basic research on reward to a biomarker for depression. Psychophysiology 52 (4), 449–459. http://dx.doi.org/10.1111/psyp.12370.
Segalowitz, S.J., Santesso, D.L., Murphy, T.I., Homan, D., Chantziantoniou, D.K., Khan, S., 2010. Retest reliability of medial frontal negativities during performance monitoring. Psychophysiology 47 (2), 260–270. http://dx.doi.org/10.1111/j.1469-8986.2009.00942.x.
Shulman, E.P., Smith, A.R., Silva, K., Icenogle, G., Duell, N., Chein, J., Steinberg, L., 2016. The dual systems model: review, reappraisal, and reaffirmation. Dev. Cogn. Neurosci. 17, 103–117. http://dx.doi.org/10.1016/j.dcn.2015.12.010.
Zisner, A., Beauchaine, T.P., 2016. Neural substrates of trait impulsivity, anhedonia, and irritability: mechanisms of heterotypic comorbidity between externalizing disorders and unipolar depression. Dev. Psychopathol. 28, 1177–1208. http://dx.doi.org/10.1017/S0954579416000754.