Is the encoding of Reward Prediction Error reliable during development?

Hanna Keren [a,*], Gang Chen [b], Brenda Benson [c], Monique Ernst [d], Ellen Leibenluft [e], Nathan A. Fox [f], Daniel S. Pine [c], Argyris Stringaris [a]

[a] Mood Brain and Development Unit, Emotion and Development Branch, National Institute of Mental Health, National Institutes of Health, 9000, Rockville Pike, Bethesda, MD, USA
[b] Scientific and Statistical Computing Core, National Institute of Mental Health, National Institutes of Health, 9000, Rockville Pike, Bethesda, MD, USA
[c] Section on Development and Affective Neuroscience, Emotion and Development Branch, National Institute of Mental Health, National Institutes of Health, 9000, Rockville Pike, Bethesda, MD, USA
[d] Section on the Neurobiology of Fear and Anxiety, National Institute of Mental Health, National Institutes of Health, 9000, Rockville Pike, Bethesda, MD, USA
[e] Section on Mood Dysregulation and Neuroscience, Emotion and Development Branch, National Institute of Mental Health, National Institutes of Health, 9000, Rockville Pike, Bethesda, MD, USA
[f] Department of Human Development and Quantitative Methodology, University of Maryland, College Park, 9000, Rockville Pike, Bethesda, MD, USA

ABSTRACT

Reward Prediction Errors (RPEs), defined as the difference between the expected and received outcomes, are integral to reinforcement learning models and play an important role in development and psychopathology. In humans, RPE encoding can be estimated using fMRI recordings, however, a basic measurement property of RPE signals, their test-retest reliability across different time scales, remains an open question. In this paper, we examine the 3-month and 3-year reliability of RPE encoding in youth (mean age at baseline = 10.6 ± 0.3 years), a period of developmental transitions in reward processing. We show that RPE encoding is differentially distributed between the positive values being encoded predominantly in the striatum and negative RPEs primarily encoded in the insula. The encoding of negative RPE values is highly reliable in the right insula, across both the long and the short time intervals. Insula reliability for RPE encoding is the most robust finding, while other regions, such as the striatum, are less consistent. Striatal reliability appeared significant as well once covarying for factors, which were possibly confounding the signal to noise ratio. By contrast, task activation during feedback in the striatum is highly reliable across both time intervals. These results demonstrate the valence-dependent differential encoding of RPE signals between the insula and striatum, and the consistency of RPE signals or lack thereof, during childhood and into adolescence. Characterizing the regions where the RPE signal in BOLD fMRI is a reliable marker is key for estimating reward-processing alterations in longitudinal designs, such as developmental or treatment studies.

Introduction

Encoding of Reward Prediction Error (RPE), the difference between the expected and received reward value, can be estimated using fMRI in humans and its alterations are thought to be involved in developmental and psychopathological processes. Yet, a basic measurement property of the RPE, its test-retest reliability, remains to be established. In this paper, we examine RPE reliability in young people (mean age at baseline = 10.6 ± 0.3 years), across 3 months and across 3 years.

The RPE is an important learning signal that helps organisms to maximize wins and minimize losses through value computations (Schultz 1998, 2006, 2013, 2016, 2017; Sutton and Barto, 1998; Rolls et al., 2008;

Diederen et al., 2016; Schultz et al., 2017). An RPE arises whenever the outcome of an action is different from what was predicted. In situations where the outcome is better than predicted, the RPE is positive and is associated with an increased likelihood of the behavior that led to the reward to re-occur. If the reward falls below what was predicted, a negative RPE occurs along with a decrease of the likelihood of repeating the same behavior. The RPE has been extensively studied in animals and found to be encoded by mesolimbic dopaminergic neurons (Olds and Milner, 1954; Corbett and Wise, 1980; Schultz et al., 1993; Bayer and Glimcher, 2005; Pan et al., 2005; Cohen et al., 2012; Averbeck and Costa, 2017).

Functional magnetic resonance imaging (fMRI) has made it possible

to localize the RPE encoding in the human brain. A recent meta-analysis of such studies indicates that the RPE is encoded in a distributed network, positive RPEs seem to be primarily represented in the striatum and negative RPEs are primarily encoded in the insula (Liu et al., 2007; Palminteri et al., 2012; Garrison et al., 2013). This has opened the way for examining the role of RPEs in sensitive stages of development, such as adolescence, and in psychopathology. Developmentally, increasing evidence suggests that reward sensitivity increases in adolescents, and, indeed, positive RPE signals in the striatum and negative RPE signals in the insula, seem to peak in adolescents compared to children or adults (Cohen et al., 2010; Somerville and Casey, 2010; Lamm et al., 2014; Smith et al., 2014; Braams et al., 2015). In psychopathology, alterations in the processing of RPEs have been proposed to be centrally involved in a range of psychiatric disorders (Murray et al., 2008; Moutoussis et al., 2015; Radua et al., 2015; Ubl et al., 2015; Schmidt et al., 2016; Rothkirch et al., 2017; White et al., 2017), including depression and schizophrenia.

Yet, despite the importance of measuring RPE in fMRI, a fundamental psychometric property remains unexamined, namely its test-retest reliability across different time scales. Test-retest reliability studies are critical for distinguishing true signal changes from other sources of measurement instability (Maitra et al., 2002; Bennett and Miller, 2010; Raemaekers et al., 2012; Herting et al., 2017). Evaluating change over time is critical for understanding developmental processes as well as psychopathology. If RPE fMRI signal is to be helpful in understanding the contribution of reward processing in these areas, then its reliability needs to be established. It is critical to understand that reliability does not represent constancy or lack of change in a measure. For example, brain activity of individuals can change over time, yet still be reliable if the rank order between those individuals in relation to the mean is maintained. This fact can also be intuited from the original formulation of the intra-class correlation (ICC) coefficient given by Fisher (1954):

$$ICC = \frac{1}{Ns^2} \sum_{n=1}^{N} (x_{n,1} - \bar{x})(x_{n,2} - \bar{x}) \quad (1)$$

where $\bar{x}$ is the pooled mean, N is the number of subjects, and the variance is given by:

$$s^2 = \frac{1}{2N} \left[\sum_{n=1}^{N} (x_{n,1} - \bar{x})^2 + \sum_{n=1}^{N} (x_{n,2} - \bar{x})^2 \right] \quad (2)$$

The difference of each individual value at each time point ($x_{n,1}$, $x_{n,2}$) is subtracted from the overall mean of the measurement occasion. It is also obvious from this formulation that reliability is inversely related to within-subject variance. When studying temporal changes, there are several sources of variance that can decrease the signal to noise ratio (SNR), such as decay in equipment calibration, or individual differences in motion parameters (Green and Swets, 1974; Horowitz and Hill, 1980; Cover and Thomas, 1991; Herting et al., 2017). Given that such noise can accumulate differentially over different time scales, it is important to estimate reliability across diverse intervals. So far, no study has addressed RPE reliability in young ages and even more so across different intervals. There have been two reports about reliability of other reward signals during adolescence (Braams et al., 2015; Vetter et al., 2017). These studies report low reliability values in mid-brain regions, where reward related signals would be typically expected. Both studies examine reliability over a single long test-retest interval of two years, which could be more influenced by cumulative errors.

In this work, we seek to establish the reliability of RPE signals across both a short (several months) and a long (several years) test-retest interval during development. We do so by using the ICC coefficient, which informs the within-subject variance relative to the total measurement variability (Bartko, 1966; Shrout and Fleiss, 1979; McGraw and Wong, 1996). For example, the popular version ICC(2,1) is defined as:

$$ICC = \frac{\sigma^2_{within\ subjects}}{\sigma^2_{within\ subjects} + \sigma^2_{between\ subjects} + \sigma^2_{error}} \quad (3)$$

As obvious from this formulation of reliability, the smaller the other sources of variability in the denominator (i.e., the between-subject variance and the measurement error), the higher (i.e., closer to 1) the within-subject reliability. We estimate the ICC using a two-way random-effects modeling approach, sometimes also referred to as a multilevel or hierarchical model, which is a powerful statistical method for estimating individual trajectories of change over time. Even though calculating the ICC measure using the ANOVA framework has been widely adopted, the application of LME methodology to ICC has several advantages in some aspects of computation where limitations are present under the ANOVA framework. Specifically, the variances for the random effects components and the residuals are directly estimated through optimizing the restricted maximum likelihood (REML) function, and thus the ICC value is computed with variance estimates instead of with their mean square counterparts under ANOVA. Therefore, in conjunction with the theoretical quantities, the estimated ICCs are nonnegative by definition. Missing data can be naturally handled in LME because parameters are estimated through the optimization of the (restricted) maximum likelihood function, where a balanced structure is not required. Moreover, incorporating confounding effects is available through adding more fixed-effects terms into the model. This LME approach for ICC has previously been implemented in the program 3dLME (Chen et al., 2013) for voxel-wise data analysis in neuroimaging. In this context, the fMRI BOLD signal change is modeled linearly via the random intercept (initial state) and slope (trajectory of change). Hence, the ICC(2,1) model is an LME case with two crossed random-effects terms. The randomization of both terms differentiates the between- and within-subject variances, enabling the estimation of within-subject reliability (Singer and Willett, 2003; Chen et al., 2013).

In this paper, we examine RPE signaling and its reliability using the "Piñata" task, a child-friendly version of the Monetary Incentive Delay (MID) task. The Piñata task has been previously shown to evoke robust reward-related fMRI BOLD activations in children and adolescents (Helfinstein et al., 2013; Lahat et al., 2016). The task elicits larger negative than positive RPE values, which occur due to "no win" outcomes in win trials. This is because in this paradigm task parameters are adjusted online to maintain a ratio of 66% of successful trials for all subjects, inducing an expectation of more positive outcomes than negative outcomes. Therefore, "no wins", when they occur, tend to induce larger RPEs relative to wins (as the latter are more expected). Subjects conducted this task in fMRI at three time points. The baseline scan (mean age 10.6 ± 0.3 years) is compared to a repeat scan following 3 ± 2.24 months and another scan following 33.6 ± 9.36 months. As a first step, we demonstrate that behavioral performance of subjects across all visits is reliable and confirm that negative RPEs predominate in this task across the three scans. For the calculation of RPE values, we follow previous studies which defined the expected value as the product of reward magnitude and the success probability (Staudinger et al., 2009; Chase et al., 2015; Ubl et al., 2015). We compare different modeling approaches for estimating the expected success probability, where each model assumes different influence of previous outcomes on the expected value. We address the question of how RPE encoding is distributed in the brain, at each one of the three scans. RPE values are used as a parametric modulator of brain activity during the reward feedback times. We test the hypothesis that negative RPEs are represented mostly in the insula while striatal regions activity is correlated to positive RPE values. We then ask whether the identified RPE signals are reliable, over three time points during development, separated by a three month and a three year test-retest interval. These results are then compared to the reliability pattern of other task activations.