NeuroImage 132 (2016) 247–260

Contents lists available at ScienceDirect

NeuroImage

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

Neuroimaging measures of error-processing: Extracting reliable signals from event-related potentials and functional magnetic resonance imaging

Vaughn R. Steele [a,b,c,*], Nathaniel E. Anderson [b], Eric D. Claus [b], Edward M. Bernat [d], Vikram Rao [b], Michal Assaf [e,f], Godfrey D. Pearlson [e,f], Vince D. Calhoun [b,c,e], Kent A. Kiehl [b,c]

[a] Neuroimaging Research Branch, National Institute of Drug Abuse, Intramural Research Program, National Institutes of Health, Baltimore, MD, USA
[b] The nonprofit Mind Research Network (MRN) & Lovelace Biomedical and Environmental Research Institute (LBERI), Albuquerque, New Mexico, USA
[c] University of New Mexico, USA
[d] University of Maryland, College Park, USA
[e] Yale University School of Medicine, New Haven, CT, USA
[f] Olin Neuropsychiatry Research Center, Institute of Living, Hartford, CT, USA

ARTICLE INFO

Article history:
Received 12 November 2015
Accepted 14 February 2016
Available online 22 February 2016

Keywords:
ERP
fMRI
Error-processing
Reliability
Time–frequency analysis
Principal component analysis
Independent component analysis

ABSTRACT

Error-related brain activity has become an increasingly important focus of cognitive neuroscience research utilizing both event-related potentials (ERPs) and functional magnetic resonance imaging (fMRI). Given the significant time and resources required to collect these data, it is important for researchers to plan their experiments such that stable estimates of error-related processes can be achieved efficiently. Reliability of error-related brain measures will vary as a function of the number of error trials and the number of participants included in the averages. Unfortunately, systematic investigations of the number of events and participants required to achieve stability in error-related processing are sparse, and none have addressed variability in sample size. Our goal here is to provide data compiled from a large sample of healthy participants (n = 180) performing a Go/NoGo task, resampled iteratively to demonstrate the relative stability of measures of error-related brain activity given a range of sample sizes and event numbers included in the averages. We examine ERP measures of error-related negativity (ERN/Ne) and error positivity (Pe), as well as event-related fMRI measures locked to False Alarms. We find that achieving stable estimates of ERP measures required four to six error trials and approximately 30 participants; fMRI measures required six to eight trials and approximately 40 participants. Fewer trials and participants were required for measures where additional data reduction techniques (i.e., principal component analysis and independent component analysis) were implemented. Ranges of reliability statistics for various sample sizes and numbers of trials are provided. We intend this to be a useful resource for those planning or evaluating ERP or fMRI investigations with tasks designed to measure error-processing.

© 2016 Elsevier Inc. All rights reserved.

Introduction

A critical aspect of designing human psychophysiological experiments is optimizing the quantity of data collected for adequately testing hypotheses. This includes both the number of participants and number of trials necessary to extract a reliable signal of interest. This is particularly important when procedures involve measurement of brain activity as these techniques require significant time and resources to collect. Prior work has often considered the numbers of trials needed to reliably measure stimulus-locked brain activity with event related potentials (ERPs) and functional magnetic resonance imaging (fMRI). Fewer investigations have addressed the stability of response-locked neural measures and the investigations that do, have focused on error-related brain activity in ERPs. The proliferation of ERP and fMRI studies examining error-related brain activity has underscored the need for definitive stability estimates for these measures. Extant studies vary widely with respect to the number of participants and the number of error trials averaged within participant. As such, ongoing research will benefit from better estimates of required numbers of trials and participants needed for stabile brain measures.

Several sources have estimated an appropriate number of trials required for stable stimulus-locked brain responses at around 20 to 50 trials. For instance, Cohen and Polich (1997) show that averages of 20 events are sufficient for attaining stability of the P300, a large ERP component related to target detection in continuous performance tasks. Stability is highly dependent upon signal to noise ratio. Luck

* Corresponding author at: National Institute on Drug Abuse, National Institutes of Health, 251 Bayview Blvd, Baltimore, MD 21224, USA.
E-mail address: vaughn.r.steele@gmail.com (V.R. Steele).

http://dx.doi.org/10.1016/j.neuroimage.2016.02.046
1053-8119/© 2016 Elsevier Inc. All rights reserved.

(2005) has suggested 30 to 60 events are sufficient for robust signals such as the P300, but also suggests that smaller components such as the N2 and P1 may require hundreds of averaged trials for reliable measures. Guidelines for fMRI studies typically recommend a minimum of 20 to 30 trials (Desmond and Glover, 2002; Huettel and McCarthy, 2001); however, these estimates are based on limited data and should be expected to vary based on the specific cognitive tasks and specifics in data collection.

In contrast to stimulus-locked brain measures, response-locked events, such as error-related brain activity, are relatively stable with far fewer trial averages (Maurer et al., in press; Meyer et al., 2013; Olvet and Hajcak, 2009; Pontifex et al., 2010; Rietdijk et al., 2014). Error processing is often assessed with ERPs and fMRI during a response inhibition task (e.g., Go/NoGo, Stroop, Stop-Signal, Flanker, Wisconsin Card Sorting Task, and Task-Switching), or any variety of speeded, continuous performance tasks likely to produce erroneous responses (for review, see (Niendam et al., 2012). The most relevant error-driven ERP components include the error negativity (Ne; Falkenstein et al., 1991) or error-related negativity (ERN; Gehring et al., 1993) and the error positivity (Pe; Falkenstein et al., 1991). These are response-locked ERP components elicited by an erroneous response to NoGo stimuli (i.e., False Alarms (FA)). The ERN/Ne is a negative deflection likely generated in the rostral cingulate zone, potentially within the caudal anterior cingulate cortex (cACC; Carter et al., 1998; Kiehl et al., 2000; Miltner et al., 2003) and peaks between 50 and 100 ms after an incorrect response (Falkenstein et al., 1990, 1991; Gehring et al., 1993; Holroyd and Coles, 2002). The ERN/Ne is believed to be associated with cognitive detection of the response error (Edwards et al., 2012; Falkenstein et al., 1991), incorrect response tendencies (Carbonnell and Falkenstein, 2006) or to reflect initial response conflict processing aimed at increasing cognitive control (Yeung et al., 2004; Yeung and Summerfield, 2012). The Pe is a positive deflection generated from at least one source within the rostral ACC (rACC; Edwards et al., 2012; van Veen and Carter, 2002) and follows the ERN/Ne, peaking between 200 and 400 ms after an incorrect response. The Pe is believed to index further error processing, conscious evaluation of the error, response strategy adjustment and/or affective assessment of the error (Endrass et al., 2007; Falkenstein et al., 1990, 1991; Leuthold and Sommer, 1999; Nieuwenhuis et al., 2001; Overbeek et al., 2005; Ullsperger et al., 2010; Yeung and Summerfield, 2012) and has been found to be negatively related with rACC activation (Edwards et al., 2012). Conscious awareness of an error is necessary for both the ERN/Ne and Pe though the ERN/Ne can be modulated by uncertainty and task parameters (Shalgi and Deouell, 2012, 2013). Successful error monitoring, as indexed by increased ERN/Ne amplitude, should lead to modulation of response strategies designed to reduce errors in the future, as indexed by reduced Pe amplitude.

The first systematic investigation of the number of trials required for a stable ERN/Ne and Pe found largely stable and reliable measures after six to eight trials among 53 young adults performing a Flanker task (Olvet and Hajcak, 2009). Rietdijk et al. (2014) found stable and internally consistent ERN/Ne and Pe with eight trials among 70 participants also performing a Flanker task. Pontifex et al. (2010) examined potential differences in reliability of error-processing across the lifespan, again with a Flanker task. The authors corroborated that six trials were sufficient for achieving stability and internal consistency of ERN/Ne and Pe in preadolescents and young adults. However, older adults may require up to eight trials to achieve the same stability. It should also be noted that while preadolescent and young adult groups had over 50 participants, the older adult group had half the participants ($n = 26$), which may have impacted the number of trials required for stability of ERPs averaged across participants. In order to extend these findings to other common error-inducing tasks, Meyer et al. (2013), examined reliability differences in ERN/Ne from three different paradigms. The authors report stable ERN, averaged across 43 participants, with six to eight trials for Flanker and Go/NoGo tasks. More errors (>20) were required to achieve acceptable reliability with the Stroop task. Each of these tasks examined averages across a set number of participants; however, reliability measures may also vary as a function of sample size (e.g., the smaller sample of older adults required more trails than the larger sample of younger adults to achieve a reliable signal). There are currently no reports that systematically examine stability of error-related activity across sample sizes and the number of events simultaneously.

The growing interest in ERP measures of error-related activity has been accompanied by an increase in MRI-based functional neuroimaging studies of these processes. While temporally less precise than scalp-recorded electrical potentials, fMRI provides more specific information about the anatomical loci supporting neural signals otherwise measured at the scalp. So, although fMRI may not temporally distinguish between the ERN/Ne and the Pe, it can measure activity in specific brain regions critical for these cognitive events with greater spatial segregation. Source localization studies from scalp-recorded ERPs suggest differentiable neural origins for the ERN/Ne and Pe. As mentioned above, the ERN/Ne arises from dorsal/caudal portions of the ACC and the Pe arises from activity in more anterior/rostral portions of the ACC (van Veen and Carter, 2002). Investigations using joint ERP and fMRI measures to evaluate neural activity error-related processes have also supported these findings, showing differentiable networks related to error processing in the rostral and caudal ACC (Edwards et al., 2012). Very little work is available that has systematically evaluated the stability of fMRI measures across sample sizes and trial numbers (Desmond and Glover, 2002; Huettel and McCarthy, 2001). None have specifically addressed stability of blood oxygenation level dependent (BOLD) activation in error-processing networks using event-related fMRI. Signal stability becomes particularly important in fMRI as inadequate power may result in a failure to identify important regions of activity and/or the mischaracterization of noise as signal of interest (see (Huettel and McCarthy, 2001). Furthermore, these parameters will vary with each cognitive task and each anatomical region of interest. Stability estimates will also depend on the experimental design (block or event-related; random or fixed effects).

Desmond and Glover (2002) provided guidelines for relative power in BOLD signal using fMRI data from passive, resting state scans (to estimate within-participant variability) and a working memory task (to estimate between-participant variability). They used these data to simulate generalizable power curves based on the number of time-points per condition and percent signal change in any given region of the brain (block design, random-effects). Results indicated that, given a percent signal change of 0.5%, a minimum of 12 participants are needed to insure 80% power at a liberal alpha of 0.05. Twice as many participants were required to achieve this power at more conservative thresholds typical of controlling for multiple tests.

Murphy and Garavan (2004) used a Go/NoGo task to examine the number of participants needed for an event-related fMRI design. They noted improvements in power and signal to noise ratio with increasing numbers of participants (from $n = 4$ to $n = 58$). The authors note that statistical power remained relatively low around $n = 20$ compared to the full sample, but low power was driven largely by type II errors rather than false-positives. For this study, the percentage of the signal evident at $n = 58$ was used to evaluate power differentials as proportions of active voxels relative to the full sample. Thus, the power achieved at $n = 58$ is assumed as an absolute ceiling. This technique does not account for effects that may remain unstable at $n = 58$, nor does it account for varying numbers of trials across participants.

The number of event trials per participant and the number of participants included in group averages are both critical variables to consider. It should be clear that there is a subtle trade-off between the number of trials averaged per participant and the number of participants involved in the study. With a larger number of participants, fewer trials may be necessary to get a reliable overall signal average. With fewer participants, a similarly reliable average might be obtained by increasing the

Download English Version:

https://daneshyari.com/en/article/6023721

Download Persian Version:

https://daneshyari.com/article/6023721

Daneshyari.com