The contribution of stimulus frequency and attention to the N2 and P3 in Go/Nogo: A multilab replication and new analyses

Giorgio Ganis[1,2], Tigan Schofield[1,2], Bradley N. Jack[3], Lisa-Marie Greenwood[3], José A Hinojosa[4,5,6], Pablo Rodríguez-Gómez[7], Darya L. Zabelina[8], Stephanie Kane[8], Kent Hubert[8], Elif Isbell[9], Faisal Mushtaq[10], Yuri G. Pavlov[11]

[1] University of Plymouth, School of Psychology, Plymouth, UK
[2] Brain Research and Imaging Centre, Plymouth, UK
[3] Research School of Psychology, Australian National University, Canberra, Australia
[4] Instituto Pluridisciplinar, Universidad Complutense de Madrid, Madrid, Spain.
[5] Dpto. Psicología Experimental, Procesos Cognitivos y Logopedia, Universidad Complutense de Madrid, Madrid, Spain.
[6] Centro de Investigación Nebrija en Cognición (CINC), Universidad Nebrija, Madrid, Spain
[7] Cardenal Cisneros, Centro de Enseñanza Superior adscrito a la Universidad Complutense de Madrid, Madrid, Spain
[8] University of Arkansas, USA
[9] University of California, Merced, USA
[10] University of Leeds, United Kingdom
[11] University of Tuebingen, Tuebingen, Germany

* Correspondence to Giorgio Ganis, University of Plymouth, giorgio.ganis@plymouth.ac.uk

Competing Interests:

The authors declare that there is no conflict of interest. Funders and employers had no role in study design or the decision to submit the work for publication.

Funding:

#EEGManyLabs is supported by a DFG (PA 4005/1-1) grant to YGP and BBSRC grant (BB/X008428/1) awarded to FM. BNJ is supported by the Australian Research Council (DE220100739).

Abstract

Numerous EEG studies have found that stimuli that require withholding a response (Nogo stimuli) elicit two scalp-recorded event-related potentials (ERPs), the frontal N2 and the frontocentral P3, that are larger than for stimuli that require a response (Go stimuli). The functional significance of these ERPs is not entirely clear, with some accounts focusing on inhibitory processes, and others focusing on a conflict monitoring explanation. Our understanding of the role of these components is further complicated by the presence of potential confounds in traditional Go/Nogo tasks. One important confound present in many studies is the lower frequency of Nogo stimuli relative to Go stimuli. Since the N2 and P3 are generally modulated by stimulus frequency and novelty, the Nogo enhancement could reflect an automatic mismatch process elicited by infrequent Nogo stimuli. A seminal study by Eimer (1993) involved matching the probability of Go/Nogo stimuli to test the contribution of stimulus frequency and manipulating spatial attention to test the explanation that an automatic stimulus mismatch process could be driving these differences. Such an explanation for the frontal N2 predicts that there should be no modulation of frontal N2 Go/Nogo effects by spatial attention. Results showed a larger N2 and P3 for Nogo than Go stimuli, indicating that stimulus frequency is not the determinant factor for these effects. Furthermore, N2 and P3 Go/Nogo effects were modulated by the attentional cue, suggesting that they are not the result of an automatic mismatch process. Subsequent research has also shown larger N2 and P3 amplitudes for Nogo stimuli when Go and Nogo stimuli are equiprobable, but the results have not been entirely consistent because of variability in the paradigms used. The purpose of the present study is to directly replicate the original study by Eimer (1993) and to test the robustness of the results to variations in the analysis pipelines. As part of the #EEGManyLabs project, 5 labs will independently replicate the original study, resulting in a multisite replication with high statistical power. Individual replications will also be integrated via a meta-analysis. Together, these studies will provide important information about the replicability of classic EEG findings and advance our understanding of the role of the frontal N2 and frontocentral P3 during inhibition tasks.

Keywords: N2, P3, Go/Nogo, spatial attention, replication, #EEGManyLabs

Introduction

Inhibitory processes are crucial for cognition and behavior because they allow the suppression of goal-irrelevant actions and information (Chambers et al., 2009). A common way to study these processes is by using variants of the Go/Nogo task. In this task, people are asked to respond (usually overtly, such as with a button press) when they perceive an imperative stimulus (hereafter referred to as “target”) denoting a Go trial, but withhold their response during Nogo trials. The key idea of Go/Nogo tasks is that each trial should elicit a prepotent Go response that needs to be inhibited on Nogo trials. In the most common type of Go/Nogo task, a single stimulus is presented on each trial, whereas in more complex versions (e.g.,S1-S2 Posner or stop-signal paradigms) targets are preceded by partially predictive cues.

Two event-related potential (ERP) components are usually modulated in Go/Nogo tasks,(1) a frontal, negative-going component between 200 and 300 ms post-stimulus (frontal N2); (2) and a frontocentral positive-going component between about 300 and 500 ms post-stimulus (frontocentral P3). Several studies have shown that these components are larger to targets in Nogo than Go trials (e.g., Albert et al., 2013; Bekker et al., 2004; Bruin & Wijers, 2002; Eimer, 1993; Enriquez-Geppert et al., 2010; Falkenstein et al., 1999; Falkenstein et al., 1995; Hong et al., 2017; Nakata et al., 2004; Pfefferbaum et al., 1985; Schroger, 1993; Smith et al., 2006, 2008). Despite being the topic of intense scientific investigation, the functional significance of these two components is not entirely clear yet, in part because a number of nearly simultaneous processes (including spatial attention and error detection) unfold during a Go/Nogo trial (e.g., Hong et al., 2017), generating ERP components that may overlap in space and time with the frontal N2 and frontocentral P3 (e.g., Albert et al., 2013).

One class of accounts of the frontal N2 effect proposes that amplitude variation in this component reflects inhibitory processes themselves (e.g., Bokura et al., 2001; Falkenstein et al., 1999; Johnstone et al., 2007; Kok, 1986). Note that inhibition processes are not restricted to motor response inhibition as frontal N2 effects occur also during Go/Nogo tasks that involve no overt response, such as covert counting (e.g., Bruin & Wijers, 2002). According to another class of accounts, the frontal N2 does not reflect inhibition per se, but captures related processes, such as conflict monitoring, that can be engaged in the absence of inhibition (e.g., Donkers & van Boxtel, 2004; Folstein & Van Petten, 2008; Groom & Cragg, 2015; Iannaccone et al., 2015; Randall & Smith, 2011; Smith et al., 2013; Yeung et al., 2004). For example, one study

(Donkers & van Boxtel, 2004) showed that the amplitude of the frontal N2 (and frontocentral P3) increased on trials when participants had to generate a response with higher force than average on a random subset of trials (i.e., the opposite of response inhibition). Yet another account holds that the frontal N2 may be an artifact of partial error trials (i.e., not resulting in an overt response, and thus included as correct trials), generating an error-related negativity (Maruo & Masaki, 2022). Similar inhibition and conflict monitoring accounts have also been used to interpret frontocentral P3 effects, with evidence mostly in favor of an inhibition interpretation (Maruo & Masaki, 2022; e.g., Randall & Smith, 2011; but see Smith et al., 2010; Wessel, 2018).

Distinguishing between these accounts empirically has proven difficult because situations requiring response inhibition usually are also associated with response conflict. For instance, to choose the correct response among a set of possible responses requires inhibiting the incorrect alternatives; at the same time, in this situation there is also conflict between these different alternatives and between them and the correct response.

Independently of the debate of whether the frontal N2 and frontocentral P3 may reflect inhibition or conflict, an important issue present in many Go/Nogo studies is that the Go stimuli are presented more frequently than Nogo stimuli (e.g., 75/25), in order to set up a prepotent Go response tendency and elicit a robust inhibitory process during the infrequent Nogo trials (Wessel, 2018). Since both the N2 and P3 are modulated by stimulus frequency and novelty, it is possible that the ERP differences between Go and Nogo trials could be due to perceptual or cognitive processes associated with infrequent stimuli (Bruin & Wijers, 2002; Daffner et al., 2000; Duncan-Johnson & Donchin, 1977; Folstein & Van Petten, 2008; Friedman et al., 2001; Tarbi et al., 2011). According to this account, Nogo enhancement of the frontal N2 could reflect an automatic mismatch process elicited by the infrequent Nogo stimuli (e.g., Daffner et al., 2000; Folstein & Van Petten, 2008; Mantysalo, 1987; Tarbi et al., 2011), similarly to the mismatch negativity (MMN) found in the auditory and visual domains (Naatanen, 1986; Tales et al., 1999). Furthermore, frequency effects could also be in part responsible for the larger frontocentral P3 to Nogo stimuli, when these are infrequent (Polich, 2007).

To test this possibility, Eimer (1993) conducted a seminal study using an S1-S2 Posner paradigm by equating the frequency of the Go and Nogo trials (50/50, Experiment 2), as well as by introducing a predictive spatial cue before the targets. The purpose of the spatial cue, an arrow pointing left or right, was to direct spatial attention toward or away from the upcoming target (corresponding to valid, 75% of the time, and invalid trials, 25% of the time, respectively). The rationale of this study is that: (i) a stimulus frequency explanation predicts that there should

not be Nogo/Go effects on the frontal N2 and/or frontocentral P3 if Go and Nogo stimuli occur with the same frequency; and (ii) an automatic stimulus mismatch explanation for the frontal N2 predicts that there should be no modulation of frontal N2 Go/Nogo effects by spatial attention.

Experiment 2 in Eimer (1993) showed a larger frontal N2 and frontocentral P3 for Nogo than Go stimuli. In addition, the frontal N2 Go/Nogo effect was modulated by the attentional cue (it was larger for valid than invalid trials). The frontocentral P3 effect was also modulated by cue validity at frontal electrode Fz, but the difference only showed marginal statistical significance. These findings suggest that stimulus frequency is not the determining factor for these Go/Nogo effects. Furthermore, they indicate that they are not due to an automatic stimulus mismatch process triggered by infrequent Nogo stimuli. Subsequent studies found that the frontal N2 and frontocentral P3 are larger for Nogo stimuli when Go and Nogo trials are equiprobable, but the results have not been entirely consistent, in part because of variability in the paradigms used (Wessel, 2018). A handful of recent studies have manipulated spatial attention in a Go/Nogo task (Hong et al., 2020; Hong et al., 2017), but the Go/Nogo paradigms used were not identical to that in Eimer (1993). For example, compared to Eimer (1993), in Hong et al (2017) participants were instructed to ignore targets on invalid trials, the cue-target delay was longer and variable and cues were different and presented at different eccentricities and polar angles.

This study by Eimer (1993) has proven to be highly influential, with several hundred citations, but there has not been a direct replication of this seminal study. The purpose of our direct, multi-lab replication of Experiment 2 from Eimer (1993) is twofold: i) to determine the replicability of the original effects using a study with higher statistical power (the original study only used 6 participants) to further validate the involvement of frontal N2 and frontocentral P3 components in inhibitory processes, and ii) to confirm and further qualify those effects with more recent analytical methods.

Based on the findings by Eimer (1993) we test 3 hypotheses. H1 predicts a larger frontal N2 at electrode Fz for Nogo than Go trials, reflecting the main effect of response assignment. H2 predicts a larger frontocentral P3 at electrode Cz for Nogo than Go trials, also related to the main effect of response assignment. Lastly, H3 predicts a larger frontal N2 effect (Nogo-Go) at electrode Fz for valid than invalid trials, indicating an interaction between cue validity and response assignment.

Methods

Statistical power and Participant Recruitment

Sample size estimation was performed through power analyses for hypotheses H1, H2, and H3. From Eimer (Eimer, 1993), the statistical results for the main effect of response assignment on the frontal N2 (H1) and frontocentral P3 (H2) were $F(1,5)=20.99$ ($p < 0.006$), and $F(1,5)=27.94$ ($p < 0.003$), respectively. Additionally, the statistics for the interaction between cue validity and response assignment on the mean frontal N2 amplitude (H3) was $F(1,5)=28.49$ ($p < 0.003$). To calculate the effect size, we used equation 13 from Lakens (2013), resulting in $\eta_p^2 = 0.808$ for the smallest of these F values. To account for the shrinkage effects typically observed in replication studies (Pavlov et al., 2021), we assumed half of this effect size. We then used G*Power (with the SPSS option) to determine the necessary sample size. This analysis indicated a sample size of 23 was required to achieve a power of 0.90 with an alpha level of 0.02. As such, each laboratory will independently collect data from a minimum of 24 participants each (an even number is required for stimulus counterbalancing, 120 participants overall).

In each replicating lab, participants will be recruited using local advertisements or online recruitment systems (e.g., SONA). Participants will be right-handed and have corrected-to-normal vision. Participants will be paid (e.g., 10£/hr in the UK and $20AUD/hr in Australia) or provided course credits for their participation. For each of the replicating labs, the study will be approved by the relevant local or national ethics committees: University of Plymouth, Faculty of Health Ethics, the ethics committee of the Universidad Complutense de Madrid, Institutional Review Board of UC Merced, Science and Medical Delegated Ethical Review Committee at the Australian National University, University of Arkansas, Institutional Review Board.

Stimuli and apparatus

Each trial (Figure 1) begins with a 200 ms presentation of a centrally located arrow (cue) pointing left or right in a random sequence (0.5 probability each). The cue points to the side of the upcoming target with a .73 (44/60) probability. A small fixation dot is presented right after the

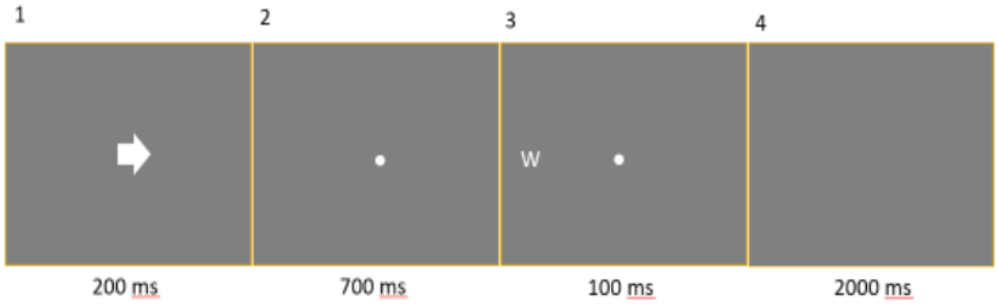

Figure 1: Trial structure. Each trial starts with an arrow pointing to the likely side of the upcoming letter, followed by a fixation dot, and the target letter (W or M, to indicate Go and Nogo trials, depending on the assignment). Participants are asked to press the key corresponding to the side of the letter on Go trials and not to respond on Nogo trials.

cue for 800 ms. 700 ms after cue offset an uppercase letter (either M or W, subtending 1 x 1 degrees) appears 6 degrees to the left or right of fixation for 100 ms. Depending on the letter assignment for that part of the study, the letter indicates a Go or a Nogo trial (0.5 probability). The interval between letter offset and onset of the next arrow is 2 s. All stimuli are presented white-on-gray on a computer display (white RGB = 255,255,255; gray RGB = 128,128,128). Note that there was no record of the actual RGB values used in the original study, and so these RGB values were obtained via personal communication with the author and represent his best estimate. The participant is seated in a dimly lit, sound-attenuated room, with response keys “1” and “0” under the left and right hand, respectively. The space bar is used to move through the instructions screens. The display is placed ~114 cm in front of the subject’s eyes and positioned so that the stimuli are shown on the participant’s horizontal straight-ahead line of sight.

Procedure

The procedure follows as closely as possible the one employed in Experiment 2 by Eimer (1993), and any known differences are reported. Participants are tested individually in an EEG laboratory and, upon their arrival in the lab, they receive a study brief and provide informed consent. Next, they are prepared for EEG recording and the EEG electrodes are attached to their face and scalp.

At the beginning of the study, participants read instructions for the task in their native language. The instructions explain that each trial begins with an arrow pointing left or right and is followed by a letter, indicating whether it is a Go or a Nogo trial, appearing on the left or on the right of a fixation dot. The task is to respond with the key corresponding to the letter side on Go trials and to withhold the response on Nogo trials. In addition to explaining the task, the instructions also emphasize the importance of keeping fixation at the center of the screen and of minimizing eye movements and blinks. After the instruction phase, participants have the opportunity to ask questions, if anything is unclear. Next, they carry out a total of 28 blocks, each composed of 60 trials and lasting 180 seconds. A brief rest period lasting 10 seconds occurs between blocks. The first 3 blocks are practice for the first letter assignment (i.e., which letter denotes Go trials), followed by 12 blocks of the corresponding experimental task. This completes the first half of the study, at which point the experimenter enters the testing room to inform the participant that the second half of the study is about to start. During this second half, the letter assignment is reversed; participants read instructions for the new letter assignment, and carry out one practice block, followed by 12 experimental blocks. Practice blocks are not included in the analyses and

letter assignment for the two sessions is counterbalanced across participants, with half of the participants tested with one letter first and the remaining half tested with the other letter first.

The whole experiment lasts approximately 2 hours, including EEG setup. Responses are acquired on a QWERTY keyboard, using the "1" and "0" keys to indicate left and right stimuli, respectively. The experiment is programmed using Presentation software (Neurobehavioral Systems, Inc., www.neurobs.com) and Psychtoolbox and translated into the local languages (English and Spanish). Additional details on the software and equipment used at each replicating lab are listed in Table 1.

Table 1

Overview of equipment and parameters used within the replicating labs

Participating lab	Operating system	Amplifier, acquisition software	Online reference, ground	Cap model, electrode type, number of cap electrodes, number of loose electrodes	Screen Type, Resolution, Refresh Rate	Stimulus Presentation, Language
University of Plymouth	Windows 10	Biosemi, ActiveTwo/Active v 9.02	CMS/DRL	Biosemi Headcap, Active (pin type for cap, flat type for the others), 32, 5 (LHE, RHE, LVE, M1, M2)	LCD, 1920x1080, 100 Hz	Presentation (NBS), English
Universidad Complutense de Madrid	Windows XP	Brain Vision actiCHamp, BrainVision Recorder	M1, FCz	Electro-Cap International, 58, 6 (LHE, RHE, LVE, RVE, M1, M2)	LCD, 1280x960, 60 Hz	Psychtoolbox, Spanish
University of Arkansas	Windows 10	Biosemi, ActiveTwo/Active v 9.02	CMS/DRL	Biosemi Headcap, Active (pin type for cap, flat type for the others), 32, 5 (LHE, RHE, LVE, M1, M2)	LCD, 1920x1080, 100 Hz	Presentation (NBS) English
Australian National University	Windows 10	Biosemi, ActiveTwo/Active v 9.02	CMS/DRL	Biosemi Headcap, Active (pin type for scalp, flat type for the others), 64, 5 (LHE, RHE, LVE, M1, M2)	LCD, 1920x1080, 60 Hz	Psychtoolbox, English
University of California, Merced	Windows 11	Brain Vision ActiCHamp Plus, BrainVision Recorder	Cz, FPz	Acticap Snap Cap, Acticap Slim, 27, 5 (LHE, RHE, RVE, M1, M2)	LED, 1920x1080, 144 Hz	Presentation (NBS), English

Electrophysiological recordings

The replicating labs will use one of the following EEG systems: (1) A Biosemi Active Two; (2) BrainAmp actiCHamp; (3) BrainAmp actiCHamp Plus. Using elastic caps, all labs will record from either 32 or 64 electrodes positioned according to the extended 10/20 EEG system

(Chatrian et al., 1985). Three or four of these 32/64 electrodes, or three additional external electrodes will be placed around the eyes to monitor the horizontal and vertical electro-oculogram (EOG), and two on the left and right mastoids to be used as reference. EOG electrodes will be attached below and above the left eye (Fp1 may be used as the matched electrode above the left eye) and on the outer canthi of the two eyes. The EEG (and EOG) data will be sampled at 500 Hz or higher (depending on the setup) with no online filters. Labs will vary in their use of electrode types (e.g., active vs. passive), online reference/ground sites (e.g., CMS/DRL, FCz, AFz, left mastoid). For details about each lab's set-up, see Table 1.

EEG preprocessing

The first data processing pipeline (referred to here as the "original pipeline") will closely replicate the one in the original study, including the following steps: activity recorded from Fz, Cz, and from the additional external electrodes will be: (i) filtered with a high-/low-pass filter of 0.1 and 70 Hz (the offline filter settings of the original study) (ii) downsampled to 150 Hz; (iii) segmented into epochs of interest (-100/+800 ms around target onset) and baseline-corrected to -100 to 0 ms prior to stimulus onset; (iv) cleaned of segments containing ocular artifacts defined as vertical EOG greater than 60 μV or horizontal EOG greater than 20 μV; (v) re-referenced to the right mastoid (M2); and (vii) trials with a response error (incorrect button press or miss for Go trials and false alarms for Nogo trials) excluded from further analyses.

Furthermore, to determine the robustness of the results, the data will also be preprocessed using a second pipeline based on more recent tools and methods (the "Current Standard" pipeline in Table 2). This pipeline includes activity recorded from all EEG sensors and consists of the following preprocessing steps carried out in EEGLAB (Delorme & Makeig, 2004): (i) downsampling to 250 Hz; (ii) trimming of EEG data segments without event codes for more than 4 seconds; (iii) removal of channels with excessive signal standard deviation (excluding eye channels and anterior frontal channels to avoid removing EOG artifacts to be corrected with ICA later) due to paroxysmal artifactual activity, using *trimOutlier*, (iv) detrending, with *removeTrend* and a 'detrendcutoff' value of 0.2 Hz in the PREP pipeline (Bigdely-Shamlo et al., 2015) (v) low-pass filtering at 40 Hz using a FIR filter (EEGLAB defaults, transition band width 10 Hz, passband edge 40 Hz, cutoff frequency (-6dB) 45 Hz); (vi) Removal of residual line noise, if still present after low-passing, (50Hz at all replicating sites, and the first 3 harmonics) using *CleanLineNoise* from the PREP pipeline, with default parameters (Bigdely-Shamlo et al., 2015); (vii) Removal of EEG data segments containing data at anterior frontal channels with absolute

amplitude larger than 300 μV or peak-to-peak amplitude larger than 600 μV; (viii) Removal of EEG data segments at any other channels (other than anterior frontal) with absolute amplitude larger than 225 μV or peak-to-peak amplitude larger than 300 μV; (ix) segmentation into epochs of interest (-200/+900 ms around the onset of the attentional cue and around the target letter), and baseline correction to -200 to 0 ms; (x) removal of epochs with horizontal EOG greater than 50 μV; (note, this step is carried out before ICA correction to ensure participants did not move their eyes just before or during the target letter, even if the horizontal EOG artifact itself could be corrected by ICA); (xi) removal of ocular or muscle artifacts using Independent Component Analysis (ICA, infomax, performed on detrended data (with EEGLAB defaults) and ICLabel (Pion-Tonachini et al., 2019) with a probability of being a muscle, heart, or channel noise artifact greater than 0.9; for ocular artifacts, removal of components for which the eye artifact weight is the largest and for which there is less than 5% brain activity weight; (xii) removal of any residual trials with vertical EOG greater than 100 μV or horizontal EOG greater than 50 μV in absolute value; (xiii) Interpolation of channels removed in step iii, if any; (xiv) re-referencing to the average of the two mastoids M1 and M2.

Outlier handling

The original study used 40% of total trials contaminated by ocular artifacts as an outlier criterion. We will use the same criterion for the replication.

Furthermore, to test the robustness of the results, we will aim to ensure good data quality by excluding participants who have more than 25% of trials rejected (i.e., 360 trials out of the 1440 trials used in total). We will calculate and report the number of included trials as well as measures of standardized measurement error (Luck et al., 2021) to summarize data quality across conditions (and across participating labs).

Quantification of the ERPs

After preprocessing, following the original analyses, the EEG will be averaged separately according to the 2x2x2x2 combination of the factors response assignment (Go/Nogo), cue validity (valid/invalid), target side (left letter/right letter), and time (first half, experimental blocks 1-12/second half, experimental blocks 13-24). This will result in 16 average waveforms per participant and electrode site. Mean amplitude values will be calculated using the original time windows: 80-130 ms (P1), 130-180 ms (N1), 250-310 ms (N2), and 300-550 ms (P3).

In addition to the direct ERP replication, we will also quantify the components leveraging the higher density recordings in our replication, by defining spatial regions of interest based on the recent literature and averaging data within these regions (Hong et al., 2017). Electrodes FCz, Cz, C1, C2, FC1, FC2 will be used to define the region of interest for the frontocentral P3. Electrodes Fz, F1, F2, will be used as region of interest for the frontal N2. If some of these electrode locations are not present in the 32 channel cap configurations used by some of the replicating labs, the largest available subset of electrodes will be used to create the regions of interest.

Statistical Analyses

Table 2:

Overview of Analyses

Analytical step	Direct replication	Robustness test 1	Robustness test 2	Robustness test 3	Robustness test 4
Pre-processing	Original	Current standard	Current standard	Current standard	Current standard
Dataset	Original electrodes and time windows	Original electrodes and time windows	Original electrodes and time windows	Regions of interest	All electrodes
Outlier handling	> 40% trials with eye artifacts	> 40% trials with eye artifacts	> 25% trials rejected	> 25% trials rejected	> 25% trials rejected
Quantification of ERPs	Mean	Mean	Mean	Mean	0-800 ms
Statistical test	ANOVA	ANOVA	HLM	HLM	HLM

The main focus of the analyses is the direct replication of the statistical methods used in the original study using repeated measures analyses of variance (ANOVAs) (Eimer, 1993). However, we will also test the robustness of these effects using hierarchical linear modeling (HLM). Table 2 reports a summary of the analyses.

(i) Direct Replication with ANOVAs

Separate repeated measures ANOVAs will be performed on mean amplitude measures for the N2 (Fz) and P3 (Cz) using response assignment (Go vs. Nogo) and cue validity (valid vs. invalid) as factors.

H1 corresponds to a main effect of response assignment in the N2 ANOVA, whereas H2 corresponds to a main effect of response assignment in the P3 ANOVA. Lastly, H3 corresponds to the interaction between cue validity and response assignment in the N2 ANOVA.

(ii) Current standard pipeline ANOVAs

The same ANOVAs used in i) will be carried out on the data preprocessed with the current standard pipeline and on the region of interest data.

(iii) Hierarchical Linear Modeling

We will also reanalyze the data by means of hierarchical linear modeling (HLM) implemented by the LIMO EEG plug-in for EEGLAB (Pernet, Chauveau et al. 2011) (see Robustness Tests 2-4 in Table 2). For these analyses, single trial data for each participant will be entered into a 1st level analysis with 16 experimental conditions resulting from the factorial combination of target location (left/right), response assignment (Go/Nogo), cue validity (valid/invalid), and time (first half/second half). The single trial data will be of 3 kinds. In order to compare the results directly with the original ones, the first kind (original electrodes dataset) will just be the single trial mean amplitudes from the original channels (Fz and Cz) for the N2, and P3 components, as described earlier. The second kind (region of interest dataset) will be the single trial mean amplitudes from the averaged regions of interest as described in the ERP quantification section. In order to have a fuller picture of the results, the third kind (all electrodes dataset) will be the single trial data from all electrodes between 0 and 800 ms post-target.

A weighted least squares solution will be used to obtain parameter estimates of each condition (Pernet, Martinez-Cancino et al. 2020). The beta estimates will then be submitted to a 2nd level analysis using a repeated measure ANOVA (generalized Hotelling T^2) with “Target Location”, “Response Assignment”, “Cue Validity”, and “Time” as factors. The contrast weights will be set up to test the difference between Go and Nogo trials (H1 and H2) and the interaction between response assignment and cue validity (H3). No correction for multiple comparisons will be used for the original electrode (as in the original study) and region of interest datasets; spatial–

temporal clustering methods implemented in LIMO will be used for multiple comparison correction for the all electrodes dataset (Maris & Oostenveld, 2007; Pernet et al., 2015).

(iv) Meta-Analysis

To address heterogeneity in EEG devices and samples between labs we will compute effect sizes for each individual site and then combine all datasets in a random-effects meta-analysis (with site as a random effect). To achieve that, first, the ANOVAs described in (i) and (ii) will be used to analyze the data from each lab separately. Next, the effect sizes of response assignment (H1 and H2) and the interaction of response assignment with cue validity (H3) found in each replicating lab will be combined. The MAJOR jamovi interface to the metafor package will be used for the meta-analysis (The Jamovi Project, 2021; Viechtbauer, 2010). The HLM analyses in (iii) will already include site as a random effect.

Assessment of the Replication and Robustness of Effects

For each hypothesis, replication success will be defined operationally as a statistically significant random-effects meta-analytic estimate (at $p < .02$) combining the results from the different laboratories, in the same direction as in the original study. Specifically, the N2 and P3 results will be considered a successful replication if the meta-analysis shows a significant main effect of response assignment (H1 and H2, Nogo > Go), and a significant interaction of cue validity and response assignment (H3, Nogo – Go effect larger for valid than invalid trials). This will confirm that stimulus frequency is not the main determinant of the N2 and P3 effects. In addition to the replication of the original study, we will also provide robustness tests by comparing these results to those obtained with current standard preprocessing methods and with the HLM analyses. If these analyses lead to the same or similar patterns of results found in (i), then it means that the original effects are not only replicable but also robust.

If some or all of the original results do not replicate, or if a different pattern of effects is found, then the direct replication will have failed and the potential reasons for this will be investigated and discussed.

Pilot data

We have collected two pilot datasets to ensure that the stimulus presentation and experimental flow control scripts worked across setups and to develop the data analysis scripts. EEG datasets were collected in Plymouth (Biosemi) and in Madrid (Brain Vision).

Data collection timeline

Starting from the date of in-principle acceptance, each laboratory will gather the data within a one-year timeframe.

Sharing of Data and Code

Pre-processing steps will be carried out using EEGLAB 2022.0 (Delorme & Makeig, 2004) implemented in MATLAB 2021b. Statistical analyses will be carried out in EEGLAB (LIMO results) and in Jamovi (ANOVAs and meta-analysis). Experimental procedures, pre-processing scripts/templates, analytical analyses and the scripts/results of the meta-analysis will be shared openly, using the Open Science Framework (OSF, https://osf.io/xawt2/?view_only=cd46ab804eac4a5e9d529b5e5eb8f740). All collected data will be made available online via https://gin.g-node.org/.

Author contributions:

Conceptualization: G.G. and Y.G.P.

Data curation: G.G., T.S., and E.I.

Formal Analysis: G.G.

Funding acquisition: Y.G.P. and F.M.

Investigation: *G.G., T.S., P.R.G., B.N.J., L.M.G., E.I., J,A,H, D,Z,, S,K,, and K.H.*

Methodology: G.G., T.S., J.A.H., and P.R.G.

Project administration: F.M., Y.G.P., and G.G.

Validation: GG, F.M., and Y.G.P.

Software: G.G. and P.R.G.

Supervision: G.G., T.S., J.A.H., B.N.J., D.Z., and E.I.

Visualization: G.G. and T.S.

Writing – original draft: G.G.

Writing – review & editing: G.G., T.S., J.A.H., Y.G.P., F.M., E.I., B.N.J., L.M.G., and D.Z.

Acknowledgements

#EEGManyLabs core team and leading lab would like to thank Prof Martin Eimer, for generously sharing critical information to ensure the fidelity and quality of this replication attempt.

References

Albert, J., Lopez-Martin, S., Hinojosa, J. A., & Carretie, L. (2013). Spatiotemporal characterization of response inhibition. *Neuroimage*, *76*, 272-281. https://doi.org/10.1016/j.neuroimage.2013.03.011

Bekker, E. M., Kenemans, J. L., & Verbaten, M. N. (2004). Electrophysiological correlates of attention, inhibition, sensitivity and bias in a continuous performance task. *Clin Neurophysiol*, *115*(9), 2001-2013. https://doi.org/10.1016/j.clinph.2004.04.008

Bigdely-Shamlo, N., Mullen, T., Kothe, C., Su, K. M., & Robbins, K. A. (2015). The PREP pipeline: standardized preprocessing for large-scale EEG analysis. *Front Neuroinform*, *9*, 16. https://doi.org/10.3389/fninf.2015.00016

Bokura, H., Yamaguchi, S., & Kobayashi, S. (2001). Electrophysiological correlates for response inhibition in a Go/NoGo task. *Clin Neurophysiol*, *112*(12), 2224-2232. https://doi.org/10.1016/s1388-2457(01)00691-5

Bruin, K. J., & Wijers, A. A. (2002). Inhibition, response mode, and stimulus probability: a comparative event-related potential study. *Clin Neurophysiol*, *113*(7), 1172-1182. https://doi.org/10.1016/s1388-2457(02)00141-4

Chambers, C. D., Garavan, H., & Bellgrove, M. A. (2009). Insights into the neural basis of response inhibition from cognitive and clinical neuroscience. *Neurosci Biobehav Rev*, *33*(5), 631-646. https://doi.org/10.1016/j.neubiorev.2008.08.016

Chatrian, G. E., Lettich, E., & Nelson, P. L. (1985). Ten Percent Electrode System for Topographic Studies of Spontaneous and Evoked EEG Activities. *American Journal of EEG Technology*, *25*(2), 83-92. https://doi.org/10.1080/00029238.1985.11080163

Daffner, K. R., Mesulam, M. M., Scinto, L. F., Calvo, V., Faust, R., & Holcomb, P. J. (2000). An electrophysiological index of stimulus unfamiliarity. *Psychophysiology*, *37*(6), 737-747. https://www.ncbi.nlm.nih.gov/pubmed/11117454

Delorme, A., & Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *J Neurosci Methods*, *134*(1), 9-21. https://doi.org/10.1016/j.jneumeth.2003.10.009

Donkers, F. C., & van Boxtel, G. J. (2004). The N2 in go/no-go tasks reflects conflict monitoring not response inhibition. *Brain Cogn*, *56*(2), 165-176. https://doi.org/10.1016/j.bandc.2004.04.005

Duncan-Johnson, C. C., & Donchin, E. (1977). On quantifying surprise: the variation of event-related potentials with subjective probability. *Psychophysiology*, *14*(5), 456-467. https://doi.org/10.1111/j.1469-8986.1977.tb01312.x

Eimer, M. (1993). Effects of attention and stimulus probability on ERPs in a Go/Nogo task. *Biol Psychol*, *35*(2), 123-138. https://doi.org/10.1016/0301-0511(93)90009-w

Enriquez-Geppert, S., Konrad, C., Pantev, C., & Huster, R. J. (2010). Conflict and inhibition differentially affect the N200/P300 complex in a combined go/nogo and stop-signal task. *Neuroimage*, *51*(2), 877-887. https://doi.org/10.1016/j.neuroimage.2010.02.043

Falkenstein, M., Hoormann, J., & Hohnsbein, J. (1999). ERP components in Go/Nogo tasks and their relation to inhibition. *Acta Psychol (Amst)*, *101*(2-3), 267-291. https://doi.org/10.1016/s0001-6918(99)00008-6

Falkenstein, M., Koshlykova, N. A., Kiroj, V. N., Hoormann, J., & Hohnsbein, J. (1995). Late ERP components in visual and auditory Go/Nogo tasks. *Electroencephalogr Clin Neurophysiol*, *96*(1), 36-43. https://doi.org/10.1016/0013-4694(94)00182-k

Folstein, J. R., & Van Petten, C. (2008). Influence of cognitive control and mismatch on the N2 component of the ERP: a review. *Psychophysiology*, *45*(1), 152-170. https://doi.org/10.1111/j.1469-8986.2007.00602.x

Friedman, D., Cycowicz, Y. M., & Gaeta, H. (2001). The novelty P3: an event-related brain potential (ERP) sign of the brain's evaluation of novelty. *Neurosci Biobehav Rev*, *25*(4), 355-373. https://doi.org/10.1016/s0149-7634(01)00019-7

Groom, M. J., & Cragg, L. (2015). Differential modulation of the N2 and P3 event-related potentials by response conflict and inhibition. *Brain Cogn*, *97*, 1-9. https://doi.org/10.1016/j.bandc.2015.04.004

Hong, X., Sun, J., Wang, J., Li, C., & Tong, S. (2020). Attention-related modulation of frontal midline theta oscillations in cingulate cortex during a spatial cueing Go/NoGo task. *Int J Psychophysiol*, *148*, 1-12. https://doi.org/10.1016/j.ijpsycho.2019.11.011

Hong, X., Wang, Y., Sun, J., Li, C., & Tong, S. (2017). Segregating Top-Down Selective Attention from Response Inhibition in a Spatial Cueing Go/NoGo Task: An ERP and

Source Localization Study. *Sci Rep*, *7*(1), 9662. https://doi.org/10.1038/s41598-017-08807-z

Iannaccone, R., Hauser, T. U., Staempfli, P., Walitza, S., Brandeis, D., & Brem, S. (2015). Conflict monitoring and error processing: new insights from simultaneous EEG-fMRI. *Neuroimage*, *105*, 395-407. https://doi.org/10.1016/j.neuroimage.2014.10.028

Johnstone, S. J., Dimoska, A., Smith, J. L., Barry, R. J., Pleffer, C. B., Chiswick, D., & Clarke, A. R. (2007). The development of stop-signal and Go/Nogo response inhibition in children aged 7-12 years: performance and event-related potential indices. *Int J Psychophysiol*, *63*(1), 25-38. https://doi.org/10.1016/j.ijpsycho.2006.07.001

Kok, A. (1986). Effects of degradation of visual stimulation on components of the event-related potential (ERP) in go/nogo reaction tasks. *Biol Psychol*, *23*(1), 21-38. https://doi.org/10.1016/0301-0511(86)90087-6

Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: a practical primer for t-tests and ANOVAs. *Front Psychol*, *4*, 863. https://doi.org/10.3389/fpsyg.2013.00863

Luck, S. J., Stewart, A. X., Simmons, A. M., & Rhemtulla, M. (2021). Standardized measurement error: A universal metric of data quality for averaged event-related potentials. *Psychophysiology*, *58*(6), e13793. https://doi.org/10.1111/psyp.13793

Mantysalo, S. (1987). N2 and P3 of the ERP to Go and Nogo stimuli: a stimulus-response association and dissociation. *Electroencephalogr Clin Neurophysiol Suppl*, *40*, 227-234. https://www.ncbi.nlm.nih.gov/pubmed/3480126

Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG- and MEG-data. *J Neurosci Methods*, *164*(1), 177-190. https://doi.org/10.1016/j.jneumeth.2007.03.024

Maruo, Y., & Masaki, H. (2022). A possibility of error-related processing contamination in the No-go N2: The effect of partial-error trials on response inhibition processing. *Eur J Neurosci*, *55*(8), 1934-1946. https://doi.org/10.1111/ejn.15658

Naatanen, R. (1986). The neural-specificity theory of visual selective attention evaluated: a commentary on Harter and Aine. *Biol Psychol*, *23*(3), 281-295. https://doi.org/10.1016/0301-0511(86)90005-0

Nakata, H., Inui, K., Nishihira, Y., Hatta, A., Sakamoto, M., Kida, T., Wasaka, T., & Kakigi, R. (2004). Effects of a go/nogo task on event-related potentials following somatosensory stimulation. *Clin Neurophysiol*, *115*(2), 361-368. https://doi.org/10.1016/j.clinph.2003.09.013

Pavlov, Y. G., Adamian, N., Appelhoff, S., Arvaneh, M., Benwell, C. S. Y., Beste, C., Bland, A. R., Bradford, D. E., Bublatzky, F., Busch, N. A., Clayson, P. E., Cruse, D., Czeszumski, A., Dreber, A., Dumas, G., Ehinger, B., Ganis, G., He, X., Hinojosa, J. A., . . . Mushtaq, F. (2021). #EEGManyLabs: Investigating the replicability of influential EEG experiments. *Cortex*, *144*, 213-229. https://doi.org/10.1016/j.cortex.2021.03.013

Pernet, C. R., Latinus, M., Nichols, T. E., & Rousselet, G. A. (2015). Cluster-based computational methods for mass univariate analyses of event-related brain potentials/fields: A simulation study. *J Neurosci Methods*, *250*, 85-93. https://doi.org/10.1016/j.jneumeth.2014.08.003

Pfefferbaum, A., Ford, J. M., Weller, B. J., & Kopell, B. S. (1985). ERPs to response production and inhibition. *Electroencephalogr Clin Neurophysiol*, *60*(5), 423-434. https://doi.org/10.1016/0013-4694(85)91017-x

Pion-Tonachini, L., Kreutz-Delgado, K., & Makeig, S. (2019). ICLabel: An automated electroencephalographic independent component classifier, dataset, and website. *Neuroimage*, *198*, 181-197. https://doi.org/10.1016/j.neuroimage.2019.05.026

Polich, J. (2007). Updating P300: an integrative theory of P3a and P3b. *Clin Neurophysiol*, *118*(10), 2128-2148. https://doi.org/10.1016/j.clinph.2007.04.019

Randall, W. M., & Smith, J. L. (2011). Conflict and inhibition in the cued-Go/NoGo task. *Clin Neurophysiol*, *122*(12), 2400-2407. https://doi.org/10.1016/j.clinph.2011.05.012

Schroger, E. (1993). Event-related potentials to auditory stimuli following transient shifts of spatial attention in a Go/Nogo task. *Biol Psychol*, *36*(3), 183-207. https://doi.org/10.1016/0301-0511(93)90017-3

Smith, J. L., Jamadar, S., Provost, A. L., & Michie, P. T. (2013). Motor and non-motor inhibition in the Go/NoGo task: an ERP and fMRI study. *Int J Psychophysiol*, *87*(3), 244-253. https://doi.org/10.1016/j.ijpsycho.2012.07.185

Smith, J. L., Johnstone, S. J., & Barry, R. J. (2006). Effects of pre-stimulus processing on subsequent events in a warned Go/NoGo paradigm: response preparation, execution and inhibition. *Int J Psychophysiol*, *61*(2), 121-133. https://doi.org/10.1016/j.ijpsycho.2005.07.013

Smith, J. L., Johnstone, S. J., & Barry, R. J. (2008). Movement-related potentials in the Go/NoGo task: the P3 reflects both cognitive and motor inhibition. *Clin Neurophysiol*, *119*(3), 704-714. https://doi.org/10.1016/j.clinph.2007.11.042

Smith, J. L., Smith, E. A., Provost, A. L., & Heathcote, A. (2010). Sequence effects support the conflict theory of N2 and P3 in the Go/NoGo task. *Int J Psychophysiol*, *75*(3), 217-226. https://doi.org/10.1016/j.ijpsycho.2009.11.002

Tales, A., Newton, P., Troscianko, T., & Butler, S. (1999). Mismatch negativity in the visual modality. *Neuroreport*, *10*(16), 3363-3367. https://doi.org/10.1097/00001756-199911080-00020

Tarbi, E. C., Sun, X., Holcomb, P. J., & Daffner, K. R. (2011). Surprise? Early visual novelty processing is not modulated by attention. *Psychophysiology*, *48*(5), 624-632. https://doi.org/10.1111/j.1469-8986.2010.01129.x

The Jamovi Project. (2021). *jamovi (version 1.6) [Computer Software]. Retrieved from https://www.jamovi.org*.

Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software*, *36*(3), 1-48.

Wessel, J. R. (2018). Prepotent motor activity and inhibitory control demands in different variants of the go/no-go paradigm. *Psychophysiology*, *55*(3). https://doi.org/10.1111/psyp.12871

Yeung, N., Botvinick, M. M., & Cohen, J. D. (2004). The neural basis of error detection: conflict monitoring and the error-related negativity. *Psychol Rev*, *111*(4), 931-959. https://doi.org/10.1037/0033-295X.111.4.931

10.1037/0033-295x.111.4.939