Placing Meta-stable States of Consciousness within the Predictive Coding Hierarchy: the Deceleration of the Accelerated Prediction Error

Amirali Shirazibeheshti[1], Jennifer Cooke[1], Srivas Chennu[2,3], Ram Adapa[4], David K. Menon[4], Seyed Ali Hojjatoleslami[1], Adrien Witon[2], Ling Li[2], Tristan Bekinschtein[5], Howard Bowman[1,6]

[1]School of Computing, University of Kent, Canterbury, CT2 7NF, UK; [2]School of Computing, University of Kent, Medway Building, Chatham Maritime, ME4 4AG, UK; [3]Division of Neurosurgery, University of Cambridge, Box 167, Cambridge Biomedical Campus, Cambridge CB2 0QQ, UK; [4]Division of Anaesthesia, Box 97, Cambridge Biomedical Campus, University of Cambridge, CB2 0QQ, UK; [5]Department of Psychology, University of Cambridge, Downing Street, Cambridge CB2 3EB, UK; [6]School of Psychology, University of Birmingham, Birmingham, B15 2TT, UK.

Address correspondence to Howard Bowman, School of Computing, University of Kent, Canterbury, CT2 7NF, UK. Tel: +44 (0)1227 823815, Fax: +44 (0)1227 762811. Email: H.Bowman@kent.ac.uk .

Abstract

While many studies have linked prediction errors and event related potentials[1] at a single processing level, few consider how these responses interact across levels. In response, we present a factorial analysis of a multi-level oddball task – the local-global task – and we explore it when participants are sedated versus recovered. We found that the local and global levels in fact interact. This is of considerable current interest, since it has recently been argued that the MEEG response evoked by the global effect corresponds to a distinct processing mode that moves beyond predictive coding. This interaction suggests that the two processing modes are not distinct. Additionally, we observed that sedation modulates this interaction, suggesting that conscious

Abrrevations: ERP (Event Related Potentials), PE (Prediction Error), GW (Global Workspace), RFT (Random Field Theory), FWE (Family Wise Error), SPM (Statistical Parametric Mapping).

awareness may not be completely restricted to a single (global) processing level.

Keywords: Levels of consciousness, local-global task, mismatch negativity, cluster-based analysis, EEG, sedation.

1 Introduction

1.1 Prediction: Predictive coding, and the associated free energy hypothesis (Friston, 2010), is one of the most prominent theories in cognitive neuroscience. Central to the theory is the notion that we are continually predicting what is likely to fall next on our sensory receptors (or more generally, the previous layer in a processing pathway). Thus, predictions are enforced top-down, while Prediction Errors (PEs) manifest bottom-up, signalling that a prediction has been infringed, see Figure 1, Panel C.

It has been argued that evoked responses reflect the confounding of expectations (Chennu et al., 2013; Friston, 2005); i.e. that they are the electrical manifestation of the brain signalling a Prediction Error (PE) (Friston & Kiebel, 2009; Friston, 2010) (strictly a precision-weighted prediction error (Feldman & Friston, 2010)). Accordingly, a number of Event Related Potential (ERP) experimental paradigms have been argued to directly engage this Prediction – Prediction Error (PE) system. A classic example of this is an auditory irregularity (oddball) task, which induces a Mismatch Negativity (MMN) response to an unexpected (so called *deviant*) tone presented within a train of (standard) stimuli (May & Tiitinen, 2010; Näätänen, Gaillard, & Mäntysalo, 1978; Näätänen, Paavilainen, Rinne, & Alho, 2007). The MMN is an increased negative deflection in the ERP, which arises soon (<200ms) after the eliciting (deviant) stimulus, with a source in auditory cortices. In addition, the

MMN is not classically viewed as involving the long distance neural exchanges argued to characterise access to a Global Workspace (GW), which putatively underlies conscious experience (Dehaene, Kerszberg, & Changeux, 1998).

Importantly, classic MMN experimental paradigms typically only consider one level of surprise. However, in real-world cognition, a single prediction is typically relevant at many different levels of processing, and indeed any resulting infringement of that prediction induces an error at many levels and many different temporal frames. For example, after the following partial sentence, “The ball rolled to John, and he ….”, the listener may predict that the sentence would be completed with a phrase such as, “kicked it”. However, hearing “picked it up”, would confound a lower level contextual prediction that John is playing football, as well as a higher level broader frame of narrative context that John is in a country that does not play rugby.

This kind of cross-level error is absolutely the norm in real-world cognition, and in this respect, the vast majority of ERP experimental work on prediction and PEs oversimplifies by only considering one level of error. Thus, to fully characterise the predictive coding framework, it is essential to understand how errors propagate in the prediction hierarchy. In particular, while experimentally manipulating different levels of surprise *in isolation* is clearly important, it can never *fully* reveal the underlying neural system and predictive architecture – interactions between levels need to be considered.

1.2 Hierarchy and Interactions: The cognitive processes that we will explore are two levels of prediction, which are called *local* and *global* in the literature. More fully, the local-global paradigm (Bekinschtein et al., 2009; Chennu et al., 2013) is a recent extension to auditory irregularity tasks that has enabled a hierarchy of predictions to

be explored. Deviance at a fine temporal resolution can, therefore, be crossed with deviance at a longer temporal resolution. The former of these is the *local effect*, with deviance at the level of tones, while the latter is the *global effect*, with deviance at the level of sequences of (five) tones (which we call *quintuples*). The local effect corresponds to the MMN, while the global effect generates a P3 component, a later (~300+ms) positive ERP deflection that is present for a globally unexpected, but not for a globally expected quintuple (Bekinschtein et al., 2009). Thus, the P3 seems to index a prediction error at a second (perhaps working memory related) level in the auditory processing pathway. This higher level is also thought to involve reverberation in long range circuits, which support engagement of a Global Workspace (GW), in a way that is typical of conscious access (Bekinschtein et al., 2009).

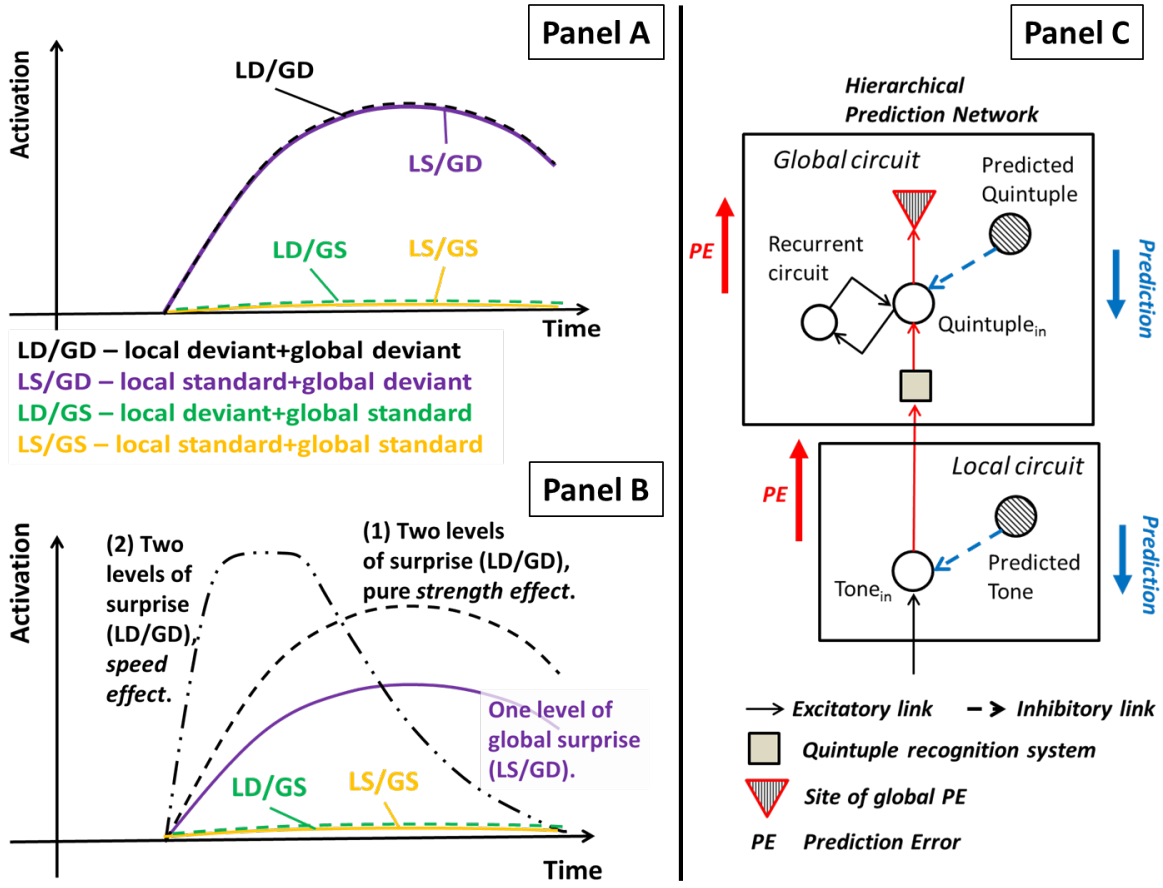


Figure 1: Caricatures of ERP effects and neural framework. (Panel A): a particular non-interaction P3 pattern, showing a main-effect of global regularity on the P3, but neither a main-effect of the local regularity, or an interaction between the two factors. (Note, the local effect/ mismatch negativity is not included in this depiction to simplify presentation.) (Panel B): potential local X global interaction pattern for the P3 response. Time series (1) shows a pure strength effect and time series (2) a speed effect, with accelerated P3. (Panel C): fragment of a putative neural architecture that could underlie the local-global effect. LD: Local Deviant; LS: Local standard; GD: Global Deviant; and GS: Global Standard.

While there has now been considerable work on the local-global paradigm (Bekinschtein et al., 2009; Chennu et al., 2013; King, Gramfort, Schurger, Naccache, & Dehaene, 2014; Wacongne et al., 2011), very little has specifically focussed on the *interaction* between local and global[2]. However, since the local-global task embeds local deviance patterns within the trains of stimuli that create global deviance, the interaction between the two levels can be explored. Accordingly, we provide a first full factorial analysis of the local-global paradigm, enabling us to characterise this interaction.

1.3 Nature of Interaction: To understand the nature of potential interactions, consider Figure 1 (Panel A), which shows candidate global effect ERP responses when there is no interaction between local and global. That is, while there is a global

2 The supplementary material of Bekinschtein et al. (2009) argued against the presence of an interaction in the local-global paradigm. However, the test they employed was non-standard and different to the classic interaction that we will consider in this article. Additionally, Wacongne et al., (2011) briefly mentioned the possible presence of an interaction, it was not characterised in depth and indeed, we will argue in the Discussion, while an interesting finding, may not be specifically revealing of the interplay of local and global processing levels.

main effect, i.e. global deviant (GD) generates a bigger response than global standard (GS)[3], the local irregularity does not differentially modulate the global level, i.e. the difference between local deviant (LD) and local standard (LS) is the same (actually zero), whatever the level of global regularity (GD or GS). In contrast, in Figure 1 (Panel B), we consider two particular ways in which levels could interact; these are *strength* and *speed*. Which of these obtains (or indeed whether both do) indicates how prediction errors aggregate across levels, as they pass up the processing pathway.

If coincidence of surprise (i.e. a global deviant that is also locally deviant) purely impacts prediction error *strength* then we should see a simple amplitude increase, without any latency change; see Figure 1 (Panel B) time series (1). However, if *speed* is also impacted, e.g. if the evoked response is accelerated, a latency-driven interaction would arise, e.g. see Figure 1 (Panel B) time series (2). Such a modulation of speed, would suggest that surprise is not always simply encoded as amplitude/strength, but, probably because of increased priority, coincidence of surprise also increases *urgency*.

1.4 Architecture: We will frame our discussion of prediction and PE in terms of an architecture such as is shown in Figure 1 (Panel C). This is a basic hierarchical prediction system that could underlie the local-global effect. It contains two levels – local and global – with each involving a prediction unit, which would be strongly active when the corresponding stimulus (a particular Tone for local circuit, Quintuple

[3] Although not shown here, another pattern that could arise, but would not constitute an interaction, would be a difference in P3 pattern between local, but not global, conditions, i.e. just a main effect of local. Although, certainly an unlikely finding, such a pattern could arise and would reflect a delayed effect of an earlier component (the mis-match negativity) on a later component (the P3). This shows that there can be temporal dependence effects, i.e. of a component at an earlier time point onto a later time point, without automatically generating an interaction.

for global circuit) is expected. Thus, a PE would be generated when an unexpected item (Tone at local, Quintuple at global) arises, since input into the circuit would not be counteracted by an active prediction unit. The resulting PEs would pass bottom-up in the processing pathway; see red links.

King, Gramfort, Schurger, Naccache, & Dehaene (2014) have argued that the brevity of the local effect means it can be modelled by a simple prediction system, such as the local circuit of Figure 1 (Panel C), but that such a system is not sufficient for the global effect. Specifically, they argue that the global effect manifests as a temporally extended invariant pattern of broad activation across the scalp. In addition, this pattern lasts for at least half a second, consistent with the notion of a Global Workspace (GW) spanning sensory, prefrontal and parietal areas (Dehaene et al., 1998), which would sustain representations through reverberating activation exchange. Indeed, in their important 2014 paper, King et al make the strong claim that local and global engage *distinct* modes. Furthermore, it has been suggested that the former of these is non-conscious, and that the latter indexes conscious processing (King et al., 2014).

We abstractly incorporate such a global level in our schematic model by assuming that a recurrent activation circuit generates the global level response; see global circuit in Figure 1 (Panel C). In this way, although certainly no more than a sketch, the framework we consider combines the predictive coding architecture with theories of brain-scale states of consciousness, especially Global Workspace Theory (Dehaene et al., 1998).

If we consider the PE generated at the global-level, see triangle in Figure 1 (Panel C), one could imagine how this form of network could generate a pure strength

pattern, i.e. time series (1) in Panel B. That is, a strong input into the global circuit, arising from a local PE, would amplify the global PE, but not generate a change in latency. In the discussion, we will describe a means by which this architecture could be extended to provide a modulation of latency, as per time series (2) in Figure 1 (Panel B).

1.5 Consciousness: In order to further characterise the interaction between local and global, we also examined the effect of pharmacologically induced changes in the level of consciousness. Such sedation studies have been undertaken in non-human primates, but these studies required animals to be awake or anaesthetised (at either moderate or deep levels of anaesthesia) and hence could not explore neural responses during wakeful sedation, which represents the critical transition between wakefulness and anaesthesia (Uhrig, Janssen, Dehaene, & Jarraya, 2016). Accordingly, we include an anaesthesia manipulation, in which sedated (unresponsiveness) and recovered (back to awareness) are compared within participants. We will also consider whether the effect of sedation is consistent across replications, i.e. trials. In particular, ERP analyses are confounded in this respect, since they investigate the *average* brain response across many trials.

One would naturally imagine that later levels of processing would be more likely to exhibit neural representations above the awareness threshold. Thus, it is important to determine whether sedation (and thus low awareness) modulates PEs early in the processing pathway, where the local effect manifests, or later, where the global effect occurs. Indeed, a strong theoretical perspective that consciousness only modulates the reverberating brain-scale state generated by the global effect would fit with a strong GW position. Most significantly, though, we will inform whether there is a clean dissociation between local and global levels, by identifying the impact of an early (putatively non-conscious) local deviant response on the later global response, which accesses neural circuitry argued to be implicated in conscious access.

Alternatively, identifying a modulatory interaction between hierarchical levels, and potentially also a further modulation by awareness level, might suggest the more complex exchanges between levels postulated by predictive coding. For example, the more advanced predictive coding formulations involving feedforward (prediction error) and feedback (prediction) exchanges, modulated by multiplicative effects of precision (Kanai, Komura, Shipp, & Friston, 2015), could generate interaction patterns.

2 Materials and Methods

2.1 Ethics statement

All participants were healthy controls, all of whom gave written informed consent. Ethical approval for testing healthy controls was provided by the Cambridgeshire Regional Ethics Committee. All clinical investigations were conducted in accordance with the *Declaration of Helsinki*.

2.2 Experimental Procedure

Bekinschtein et al. (2009) devised an auditory experiment to explore the relationship between consciousness and ERP components like the P3, and MMN. The resulting local-global experiment, see Figure 2, is composed of quintuples (sequences of five tones), where the first four tones are identical in frequency but the last can be either the same as the four preceding (locally standard) or change to a higher or lower pitch (locally deviant). In addition to its local regularity, each of the last quintuples in Figure 2 is itself either *globally standard*, where the quintuple appears frequently, or *globally deviant*, where the quintuple appears rarely throughout the prior context. Depending on how the local and global regularities are manipulated, different trial types can be defined. We denote these trial types LDGD (Local Deviant Global Deviant), LSGS (Local Standard Global Standard), LDGS (Local Deviant Global Standard) and LSGD (Local Standard Global Deviant). Additionally, the sedation manipulation, identified with a prefixed R or S, for Recovered or Sedated, gives a full

set of eight conditions: RLDGD, RLSGS, RLDGS, RLSGD, SLDGD, SLSGS, SLDGS and SLSGD.

The global standard trials arise ~80% of the times, whereas the global deviant trials arise ~20% of the times. The global deviant condition comprises the trials of conditions (1) and (4), while the global standard comprises the trials of conditions (2) and (3). The *global effect* is simply the average of global deviant trials minus the average of global standard trials. By the same token, the average of the local deviant trials, i.e. of conditions (1) and (3), minus the average of the local standard trials, i.e. of conditions (2) and (4), comprises the *local effect*. The local-by-global interaction is then calculated as the difference of differences from these four conditions (e.g. (LDGD-LSGD)-(LDGS-LSGS)). 20-30 global standard quintuples were repeated before the occurrence of the globally deviant quintuple to establish the global regularity. The local-global task was presented when participants were sedated and also once they had recovered from sedation, 20 minutes later. The participants were requested to actively count the occurrence of the global deviant trials, while high density EEG was recorded.

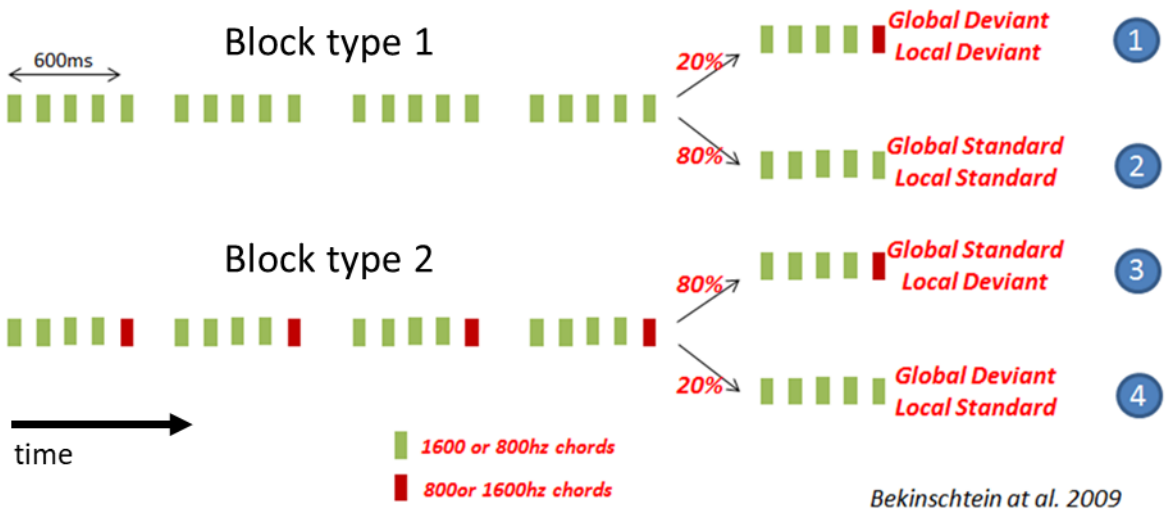


Figure 2. The local-global task: (condition-1) local deviant-global deviant (LDGD); (condition-2) local standard-global standard (LSGS); (condition-3) local deviant-global standard (LDGS); and (condition-4) local standard-global deviant (LSGD). Quintuples (i.e. five tone sequences) were either locally standard (all tones the same) or locally deviant (5th tone, shown in red, different). Blocks were defined by the quintuple type (locally standard or locally deviant) that was frequent (i.e. was globally standard). Local standard quintuples are globally standard in block type 1, while locally deviant quintuples are globally standard in block type 2. After a series of quintuples to set the global context for a block, global deviants (condition (1) for block type 1 and (4) for block type 2) occur 20% of the time, and global standards (condition (2) for block type 1 and (3) for block type 2) the other 80% of the time. The duration of each tone is 50ms and the time interval between the onset of two successive tones is 150ms presented via headphones with an intensity of 70dB. All tones were generated with 7ms rise and fall times. Participants were requested to actively count the occurrence of global deviants, while high density EEG was collected.

2.3 Drug Administration

Participants were sedated with the anaesthetic agent propofol. Each experimental run began with an awake baseline period lasting 25-30 minutes followed by a target-controlled infusion of Propofol (Marsh, White, Morton, & Kenny, 1991), administered via a computerized syringe driver (Alaris Asena PK, Carefusion, Berkshire, UK). With such a system, the anesthesiologist inputs the desired ("target") plasma concentration, and the system then determines the required infusion rates to achieve and maintain the target concentration (using the patient's characteristics, which are covariates of the pharmacokinetic model). The Marsh model is routinely used in clinical practice to control propofol infusions for general anesthesia and for sedation. Two blood plasma levels were targeted – 0.6μg/ml (mild sedation), 1.2μg/ml (moderate sedation), and recovery from sedation. A period of 10 minutes was allowed for equilibration of calculated and actual plasma propofol concentrations before cognitive tests were commenced. Blood samples (for plasma propofol levels) were drawn towards the end of each titration period. In total, 5 blood samples were drawn during the study, at the beginning and end of each target plasma level (2 for mild and 2 for moderate sedation) and one after the end of the theoretical washout (during the recovery period). The state of mild sedation was aimed to engender a relaxed but still responsive behavioural state. Computer simulations with the TIVATrainer pharmacokinetic simulation software revealed that plasma concentration of propofol would approach zero in 15 minutes leading to behavioural recovery; hence behavioural assessment was recommenced 20 minutes after cessation of sedation and the final blood sample.

These data were acquired as part of a larger experimental protocol, which precluded acquisition of responses to the relevant stimuli in the awake state – these were only obtained while the participant was sedated (Sedation), and then when the participant had behaviourally recovered from sedation (Recovered).

2.4 EEG Recording and Pre-processing

129-channel high-density EEG data were collected at each level of sedation, measured in microvolts (uV), sampled at 250Hz and referenced to the vertex, using the Net Amps 300 amplifier (Electrical Geodesics Inc., Oregon, USA). Participants had their eyes closed during data collection. From the continuous EEG data, we rejected voltages exceeding ±200uV, including transient spikes that exceeded ±100 uV. In addition, electrooculogram activity exceeding ±70uV was rejected. Trials were then segmented from -200 ms to 1300 ms relative to the onset of the first sound, providing full quintuple segments. Bad channels were removed and interpolated if they crossed a threshold of 200uV. Trials with greater than 25 bad channels were rejected. After artefact rejection, the remaining full quintuple trials were referenced to average, band-pass filtered from 0.5Hz to 20Hz, and corrected in EEGLAB for the baseline duration of 200ms (from 400ms to 600ms, just before the onset of the fifth tone). Consistent with this baselining, the onset of the fifth tone becomes the zero point, i.e. 600ms in quintuple length segments becomes 0ms for our remaining analyses and figures. Additionally, full quintuple trials were then segmented in EEGLAB into intervals (for analysis) from 50ms (ending of the fifth tone in the quintuple) to the end of the full quintuple trial, before converting to SPM EEG-image format. We call the resulting segments *trials.* For display purposes, in SPM, the scalp topographies and statistics are presented from 100ms to 580ms relative to the onset

of the fifth tone in a quintuple, at each 32ms time bin. 22 participants were originally tested, but 4 were lost due to technical issues.

2.5 EEG Analysis

2.5.1 ERP Analysis – SPM for EEG

Before further analysis, channels near the neck and eyes, 36 out of the 129 channels, were discarded. This was done firstly, since these channels are often sources of muscle and eye-movement artefacts, and secondly, because of confounding edge effects that arise at side channels due to the average referencing (Chennu et al., 2013). A 2 by 2 by 2 repeated measures ANOVA was set up in SPM, both at the participant and group level. That is, the full factorial ANOVA analysis consisted of three factors, each factor with two levels. The factors were Sedation (sedated vs recovered states), Global regularity (standard vs deviant) and Local regularity (standard vs. deviant). The average effect of each participant in each condition (i.e. beta-images) at the participant level formed the subject input matrix to the group level analysis.

Additionally, we explored the consequences of including a number of covariates of no interest in our design. These covariates encoded cross-participant variability in behavioural responses (which in the Recovered case could be used as a proxy for recovery from sedation) and whether the participant reached mild or moderate sedation, as determined by blood plasma level – see section 2.3. None of the combinations of covariates we considered changed the results of our SPM analysis substantively. This suggests that none of these measures encode influences beyond those already accommodated in our regression model. As a result, for simplicity, we present an analysis without covariates of no interest.

Using SPM12, we present one of the first uses of a cluster extent ERP analysis. There has been considerable recent debate concerning the appropriate cluster-forming threshold to use with RFT, e.g. (Eklund, Nichols, & Knutsson, 2016). Accordingly, we have taken the safe approach of enforcing the default fMRI cluster forming threshold of 0.001 and FWE rate of 0.05 (Friston et al., 1996; Woo, Krishnan, & Wager, 2014) in our ERP analysis. This choice reduces the probability of creating thresholded random fields with multiple peaks in the same cluster, the situation in which the Euler Characteristic can misestimate the probability of a cluster size under the null hypothesis (Flandin & Friston, 2016).

In our analysis, we kept in all trials that passed artefact rejection. Section S6 of the Supplementary Material elaborates on this decision, given that conditions have different trial counts.

2.5.2 Single Trial Analysis

A further question we explore is whether the variability in the evoked response changes with sedation. For example, it could be that awareness (tonically) waxes and wanes during sedation (especially when that sedation is only partial), in the manner that wakefulness can vary during the transition to sleep (Bareham, Manly, Pustovaya, Scott, & Bekinschtein, 2014). We call this the *drifting tonic alertness* hypothesis. We will in fact find (see for example, bottom Panel of Figure 4) that sedated ERP responses are, indeed, generally reduced in amplitude. Such a reduction in the “across trials” average (i.e. the ERP), could arise from an “across-the-board” decreased single trial amplitude or alternatively from drifting tonic alertness, which might cause a number of trials to be completely missed when sedated.

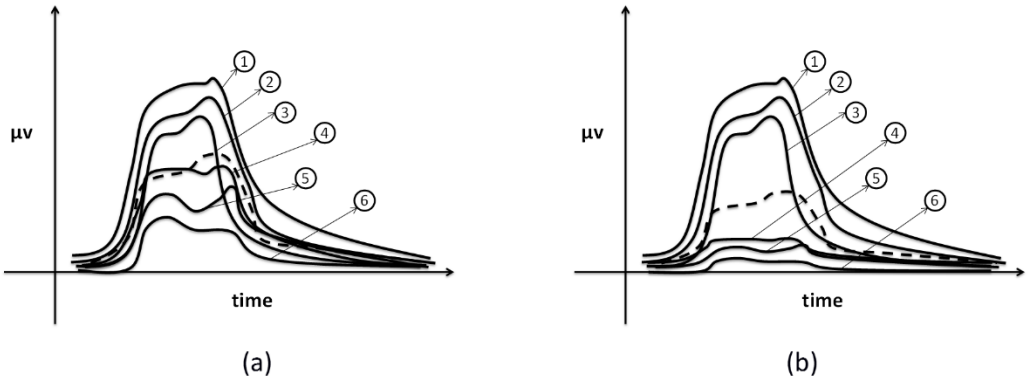


Figure 3. Illustration of variation across replications (i.e. single trials) in two conditions. The solid lines represent the (single trial) replications of a component (of which there are six, in each condition, with each replication labelled with a number) and the dashed lines (one in each condition) show the average across the replications (i.e. they are the ERPs). In our context, the condition in (a) might represent the situation when participants are recovered, and amplitude variation is low, while the condition in (b) might represent when participants are sedated and alertness is waxing and waning. That is, in (a) there is more graded variation in amplitude across single trials than in (b).

Figure 3 illustrates this, with the simplifying assumption of no temporal jitter across trials. The solid lines represent the single trial EEGs and the dashed lines the ERP responses. While the amplitudes of trials 1 to 3 are similar in the two conditions (Figure 3.a and Figure 3.b), the amplitudes of trials 4 to 6 in the second condition (Figure 3.b) are much lower than in the first condition (Figure 3.a). This causes a lower average effect across trials (dashed curve) in the second condition relative to the first. In other words, the lower ERP amplitude in the second condition is due to higher variation in amplitude, not an “across the board” reduced single trial component amplitude. (Importantly, this pattern is distinct from the known property

that variability increases with mean amplitude, as, for example, is encapsulated in the Fano factor (Dayan & Abbott, 2001), since our hypothesis is driven by the difference in how graded the variability is across trials.)

To address this issue, we will determine the amplitude of the P3b (the component we will be interested in) elicited by global deviance on each single trial, and then calculate the standard deviation of the resulting distribution of amplitudes.

As will become clear, there is variability in latency of the P3 across conditions. As a result, analysis windows could be placed in a number of ways. Accordingly, we will consider two different strategies for placing analysis windows: Same Window Across Conditions and Varying Window Across Conditions, which we discuss in turn shortly.

2.5.2.1 Average amplitude analysis - For both the window placements we use, the same basic analysis was applied. Specifically, the average amplitude of each trial over the identified window associated with global deviance was calculated and then the standard deviation of these average amplitudes was calculated for each participant and condition. Finally, a 2x2x2 (local x global x sedation) ANOVA was conducted with 18 participants, 8 conditions, and with the standard deviation (of amplitude across trials) as the dependent measure. This design was applied for both channels Cz and Pz (see ERPs in Figure 6, second panel down (1st containing ERPs)), where the effect of sedation on the P3 amplitude is most evident.

For the 3-way full factorial ANOVA, the sedation-global interaction is the most important, since it enables us to find the effect of sedation on the difference in across-trial variation between global deviant and global standard. To put this in another way, we are interested to know if the reduced amplitude, which we will observe of the P3 when sedated relative to recovered is consistent across trials.

To give as much opportunity as we can for difference in the standard deviation of amplitude to be observed, we perform multiple analyses, each with different window placement, since there are a number of different plausible ways in which windows could be placed. Importantly, if we fail to find a difference in standard deviation in all these analyses, then there is little evidence to reject the null hypothesis. This is because we will have given the alternative hypothesis many ways to be demonstrated. Notice also that “fishing” criticisms of selective window placement, e.g. (Brooks, Zoumpoulaki, & Bowman, 2016; Kilner, 2013), (which is related to double dipping as discussed in the fMRI literature (Kriegeskorte, Simmons, Bellgowan, & Baker, 2009)) do not apply for this analysis, since we will be seeking to demonstrate a null effect. Put in other words, if, in this situation, we do *not* find evidence to affirm the alternative hypothesis, then we can be even more confident that we should not be rejecting the null hypothesis.

2.5.2.2 Same Window across Conditions - In our first approach, for all conditions a window was taken from 200ms up to the end of the analysed segment, where the P3 presents itself in the ERPs; see ERPs at channels Cz, and Pz in Figure 4, lower part of the middle panel, marked G.

2.5.2.3 Varying Window across Conditions - Fixing the window across conditions (with the 200ms to end window), will mean that the optimal window is not selected for each condition. This is especially true when there are latency differences across conditions, which there are here. So, to provide further evidence, as discussed earlier, a second approach was applied, in which the average amplitude window for RLDGD (Recovered Local Deviant Global Deviant) was set from 200ms to 400ms, where the P3 appears as a sharp but short response; see ERPs at channels Cz, and Pz in Figure 6, panel of 2nd row down. The same time window was applied over all

the LD conditions, including SLDGD (Sedated Local Deviant Global Deviant), RLDGS (Recovered Local Deviant Global Standard), and SLDGS (Sedated Local Deviant Global Standard). The GS does not generate a strong P3, c.f. GS ERPs in Figure 6, panel of 2nd row down, or the relevant ERPs for the local-global interaction (LxG) in Figure 5, lower part of top panel, thus we have applied a window from the key comparator condition that contains a large P3. But importantly, the average amplitude window for RLSGD was set from 250ms to the end of the time region, where the P3 appears as a slow but long response; see ERPs at channels Cz, and Pz in figure 6, panel of 2nd row down. The same time window was applied over all the LS conditions, including SLSGD, RLSGS, and SLSGS.

2.5.2.4 Bayesian Analysis - The classical (ANOVA) statistical test for average amplitude variation will be followed by a Bayesian analysis to find evidence in favour of the null hypothesis that there is no difference in standard deviation across trials when sedated versus recovered. Hence, the Bayes factor (BF) will be calculated using a Bayes calculator developed by (Dienes, 2014), and we will show that there is evidence to support the null. The Bayes factor (BF) is the ratio of likelihoods for the alternative to the null hypothesis. A bigger ratio corresponds to higher evidence for the alternative, and a smaller ratio to higher evidence for the null. The BF is generally divided into three areas (Dienes, 2014). If it is more than 3 ($B>3$) or less than 1/3 ($B<1/3$), there is substantial evidence for the alternative or null, respectively. The test is not sensitive for a BF between 1/3 and 3 ($1/3<B<3$).

2.5.2.5 Prior distribution - To perform Bayesian analysis, we use uniform priors. However, these require an estimate for the upper and lower limits of these uniform priors (Dienes, 2014). One way to do this is to use constraints within the dataset itself to determine the plausible parameter range (so called, empirical priors). Since

clear precedents in the literature are not available, to find the range of priors, we apply the limit of a 95% confidence interval for each condition (Dienes, 2014). The sedation-global interaction (SxG) is (RGD - RGS) – (SGD - SGS), where each term of the interaction is collapsed across the local effect; e.g. RGD = (RLDGD+RLSGD)/2. Therefore, the interaction can be written as follows:

$$S\times G = \left(\frac{RLDGD + RLSGD}{2} - \frac{RLDGS + RLSGS}{2}\right) - \left(\frac{SLDGD + SLSGD}{2} - \frac{SLDGS + SLSGS}{2}\right) \quad (1)$$

The interaction is maximal when the first and last terms (RGD and SGS) are maximal and the second and third (RGS and SGD) are minimal, which will be obtained when RLDGD, RLSGD, SLDGS, and SLSGS take the upper limits of their confidence intervals, and RLDGS, RLSGS, SLDGD, and SLSGD take the lower limits of theirs'. By the same token, the interaction is minimal when RLDGD, RLSGD, SLDGS, and SLSGS take the lower limits of their confidence intervals, and RLDGS, RLSGS, SLDGD, and SLSGD take the upper limits of theirs'. The maximum of the absolute values of the upper and lower limits determines the priors' upper and lower limits. e.g. if the maximum of SxG is 3.5 and its minimum is -4, the maximum of the absolute values of the upper and lower limits is 4. So, the prior range will be -4 to 4.

2.5.2.6 Likelihood distribution: The BF calculator of (Dienes, 2014) assumes a normal distribution for the sampling distribution of the parameter estimate. However, if degrees of freedom (df) are less than 30, the distribution standard error needs to be corrected by a factor (1+20/(df x df)), which is the case in our analysis (18 participants, df =17).

To calculate the mean difference of the likelihood distribution for the sedation-global interaction, i.e. mean of (RGD - RGS) – (SGD - SGS), we need to calculate the average effect size of each of the eight conditions in equation (1). Furthermore, the

standard deviation of the normal distribution can be estimated using the standard error of the denominator of a t-test. Having F-values from the full factorial analysis of average amplitude, t-values are simply the square root of the F-values; i.e. $t = \sqrt{F}$. Finally, the standard error is the division of the mean difference (MD) by the t-value; i.e. MD/t.

Bayes factors: Having the distribution of likelihoods together with the limits of the priors, the BF is calculated as the ratio of the probability of the observed data under the alternative hypothesis, to the probability of that data under the null.

3. Results

This section consists of two sets of analyses, presented in two subsections. The first presents the results of a full factorial SPM for EEG analysis, and the second a single-trial analysis. We only present ERP time-series at key significant electrodes; the time-course at other electrodes can be seen in (Shirazibeheshti, 2015). Additionally, text and a table summarising the statistical findings of our cluster-level analysis are presented in the Supplementary Material; see subsection S3, “Summary of Cluster-level Analysis” and table S.T1.

3.1 SPM for EEG Full Factorial Analysis of ERPs

In the following ERP depictions, e.g. Figures 4 to 6, labelling of time axes is relative to the onset of the fifth tone in the relevant quintuple.

3.1.1 The local effect

As previously shown (Bekinschtein et al., 2009), the local effect (local deviant minus local standard) is typically divided into two phases. The first phase is the classic

mismatch negativity (MMN), and the second a positive (P3a) rebound. In our data, the former occurs frontally from 100 to 132ms, while the latter occurs centrally from 196 to around 356ms; these phases are shown for the local effect in Figure 4, top panel, marked L. Cluster level analysis for local effect, see Figure 4, second row of scalp topographies in top panel, reveals that the effect is significant at both negative and positive phases, and that it extends (frontally) over an even longer time window into a third phase. However, clusters are smaller in size and weaker at later times. The local effect ERPs are presented at Fz and Cz in Figure 4, lower part of top panel.

3.1.2 The global effect

The global effect (global deviant minus global standard) is posterior positive and significant during a time window from (around) 260ms post stimulus onset up to the end of the time-series (illustrated in Figure 4, middle panel, marked G). However, initially the effect is weak, only becoming substantial around 356ms post stimulus onset. This global effect corresponds to a P3b. ERPs for this effect are presented in Figure 4, the lowest part of middle panel, at central and posterior channels; i.e. Cz and Pz. The effect is a significant positive response at Cz and even stronger at Pz.

3.1.3 The sedation effect

For the main effect of sedation, Figure 4, bottom panel (marked S), the effect (recovered minus sedated) is significant in a time window from 100 to 132ms post stimulus onset, followed by a second significant time window from 260 to 388ms. It starts as a negative anterior effect, which changes through time to a positive posterior effect.

The ERPs at channels Fz and Pz are presented in Figure 4, bottom row of the lower panel. The effect is negative and early at channel Fz. Following this there is a later positive response at Pz. Broadly speaking, the size of the EEG effect, in the sense of distance from zero, is always greater for recovered than sedated, i.e. when the ERPs are negative, recovered is smaller than sedated, and when they are positive, recovered is bigger than sedated.

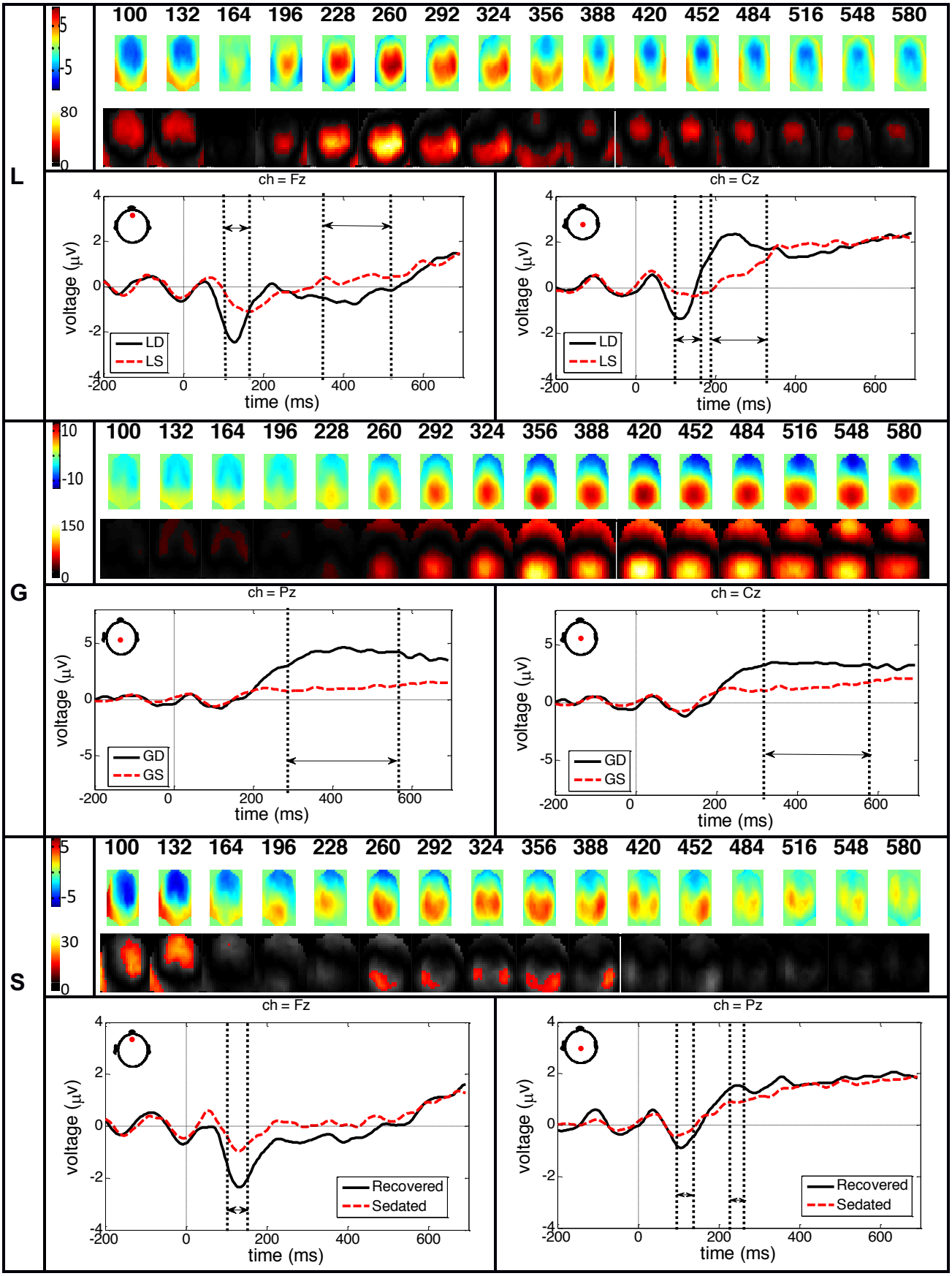


Figure 4. The three *main* effects: the local effect (L=LD-LS), global effect (G=GD-GS), and sedation effect (S=Recovered-Sedated). Sequences of scalp

maps through time are shown, with front of head at the top of each scalp. T-values and thresholded F-values (output of Random Field Theory (RFT)) for each effect is presented at the first and second rows of each panel, respectively. Side channels are removed, cluster forming threshold is set at 0.001, and Family-wise Error (FWE) rate correction at 0.05. The corresponding ERPs for each effect are also depicted. The double direction horizontal arrows show the region where the effects are significant. The onset of the fifth tone is at zero, the time point from where the ERP responses are analysed to find the effect. The window from -200 to 0ms is the baseline region.

3.1.4 The local-global interaction

The significant local and global effects in Figure 4, the top and middle panels, partially overlap during the second phase of the local effect and subsequently where the local effect is even significant over the P3 (global effect) time window. This raises the possibility that there may be an interaction between the two effects. The top panel of Figure 5, marked LxG, presents this interaction. The interaction pattern starts with a significant positive centro-posterior cluster emerging at 260ms post stimulus onset, which then switches polarity to a negative centro-anterior cluster, which becomes substantial around 452ms and continues up until the end of the time-series.

The local-global interaction ERPs in Figure 5, lower part of top panel, suggest an *accelerated* (in time) global effect in LDGD compared to LSGD, while the global standard pattern changes little between conditions. The LSGD response onsets later in time and is longer in duration, but interestingly, is not much lower in amplitude. The difference between LDGD and LSGD is much larger than the difference

between LDGS and LSGS (see the ERPs at channel Cz in Figure 5, lower part of top panel, for example). This suggests that *the global deviance response is modulated by the level of the local effect; while global standard responses are influenced very little by the local level*. To explain the interaction in more detail, we have also given an example of calculation of the interaction in the Supplementary Material; see Figure S1.

In fact, when considering the interaction at a single point, the line of causation is unclear, i.e. is it the local effect that is modulating the global or the global modulating the local? We return to this issue of interpretation in the discussion.

3.1.5 The sedation-local interaction

The sedation-local interaction is presented in Figure 5, middle panel, marked SxL. It is clear that sedation and recovery differ within the time windows of the positive (P3a) rebound of the local effect from 228 to 292ms. The presence of a sedation-local interaction suggests that the local effect is not isolated from the effects of sedation, and thus is not qualitatively different from the global effect by virtue of the impact of sedation on it. The SxL interaction is not significant in the first negative phase of the local effect (i.e. the N1); so, the very early local effect (the MMN) is not found to be significantly modulated by sedation, although there is a trend towards significance.

To explore this further, the lower part of the middle panel in Figure 5 shows ERPs of the sedation-local interaction at two electrodes. It is evident that when sedated, the size of the local effect (SLD-SLS) is reduced and possibly slowed in its second (positive rebound) phase in comparison to when recovered; i.e. RLD-RLS. In other

words, the RLD is stronger and (perhaps) accelerated compared to the other three conditions.

3.1.6 The sedation-global interaction

The bottom of Figure 5, marked SxG, shows the presence of a sedation-global interaction. This is an indication that the global effect can be modulated by sedation, which is consistent with Bekinschtein et al.'s (2009) proposal.

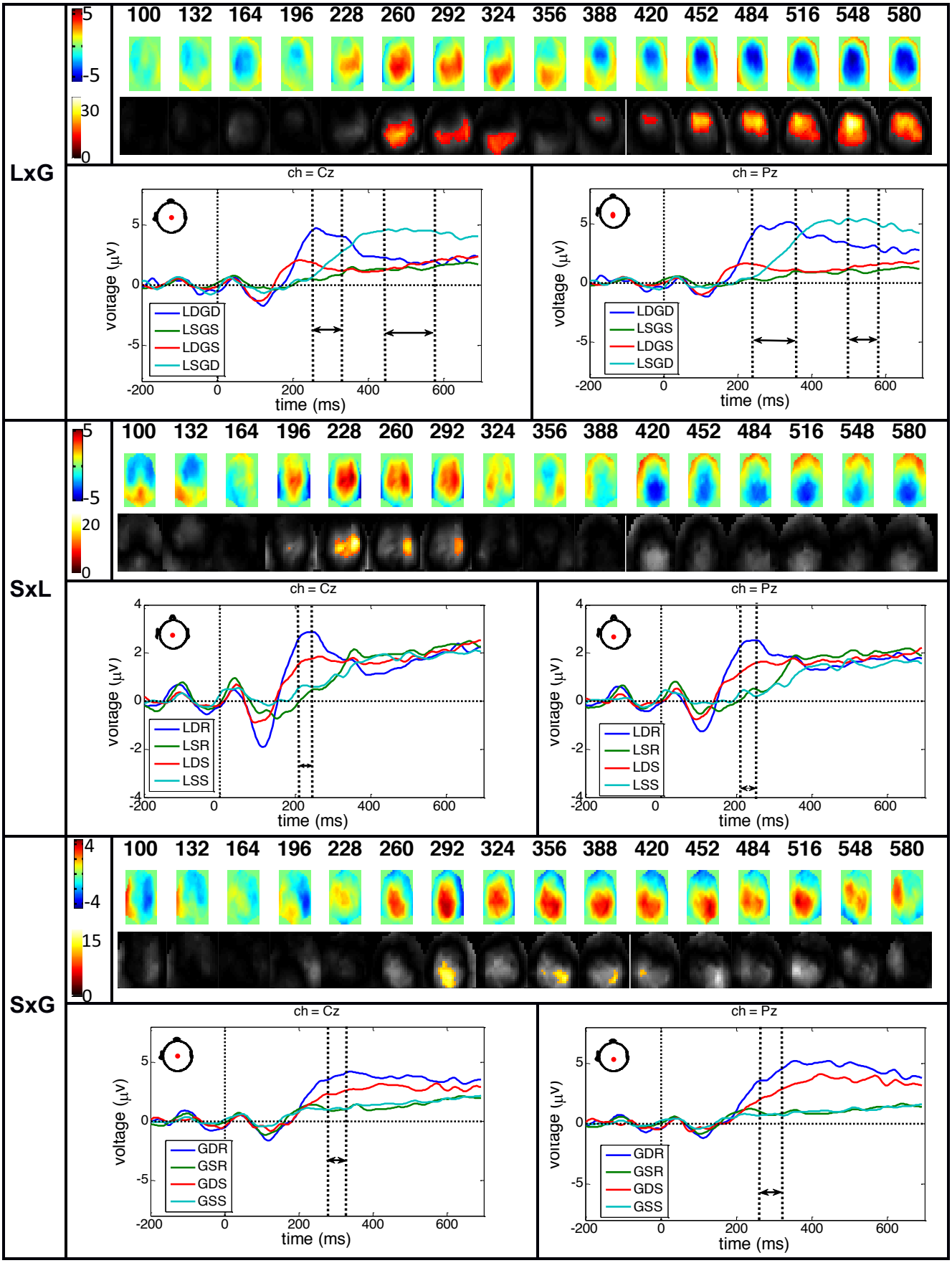

LxG
100 132 164 196 228 260 292 324 356 388 420 452 484 516 548 580
5
-5
30
0
ch = Cz
ch = Pz
voltage (μV)
time (ms)
LDGD
LSGS
LDGS
LSGD
SxL
20
LDR
LSR
LDS
LSS
SxG
4
-4
15
GDR
GSR
GDS
GSS

Figure 5. The local-global (LxG), sedation-local (SxL), and sedation-global (SxG) interactions. Sequences of scalp maps through time are shown, with front of head at the top of each scalp. T-values, and thresholded F-values (output of the Random Field Theory (RFT)) for each effect is presented at the first and second rows of each panel, respectively. Side channels are removed, cluster forming threshold is set at 0.001, and Family-wise Error (FWE) rate correction at 0.05. The corresponding ERPs for each effect are also depicted. The double direction horizontal arrows show the region where the interactions are significant. The onset of the fifth tone is at zero, the time point from where the ERP responses are analysed to find the effect. The window from -200 to 0ms is the baseline region. LD (local deviant), LS (local standard), GD (global deviant), GS (global standard), R (recovered), S (sedated).

The ERPs in the lower part of bottom panel in Figure 5 for the sedation-global interaction also illustrate the decreased size (amplitude) of the global effect when sedated vs. recovered; ie. (RGD-RGS)>(SGD-SGS). The findings are also suggestive of a delayed onset for the global effect, especially more posteriorly when sedated.

3.1.7 The local-global-sedation (3 way) interaction:

The 3-way interaction in Figure 6, top panel, shows how sedation modulates the local-global interaction. The 3-way interaction we observe is a centro-posterior region of negativity, occurring around 420ms post fifth stimulus. Figure 6, second panel down, shows the time-series for all conditions at Cz and Pz. The period of 3-

way significance is late, appearing around 420ms (c.f. Supplementary Material, Figure S2 for finer resolution F-maps).

The interaction plots, middle column of the bottom panel of Figure 6, show the effect most clearly. These plots depict the amplitudes of ERPs at the point of the grey double arrows. They show the two way (LxG) interaction separately for Recovered (top row of lower panel) and when Sedated (bottom row of lower panel). One can view the 3-way interaction as arising from a change in the LxG interaction between Recovered (top row) and Sedated (lower row). Since there is little difference between LDGS and LSGS conditions (in fact, all four GS conditions are very similar), the interaction can be seen to be driven by the amplitude of the LSGD condition relative to LDGD, which is considerably larger when Recovered (see, X_R) than when Sedated (see, X_S).

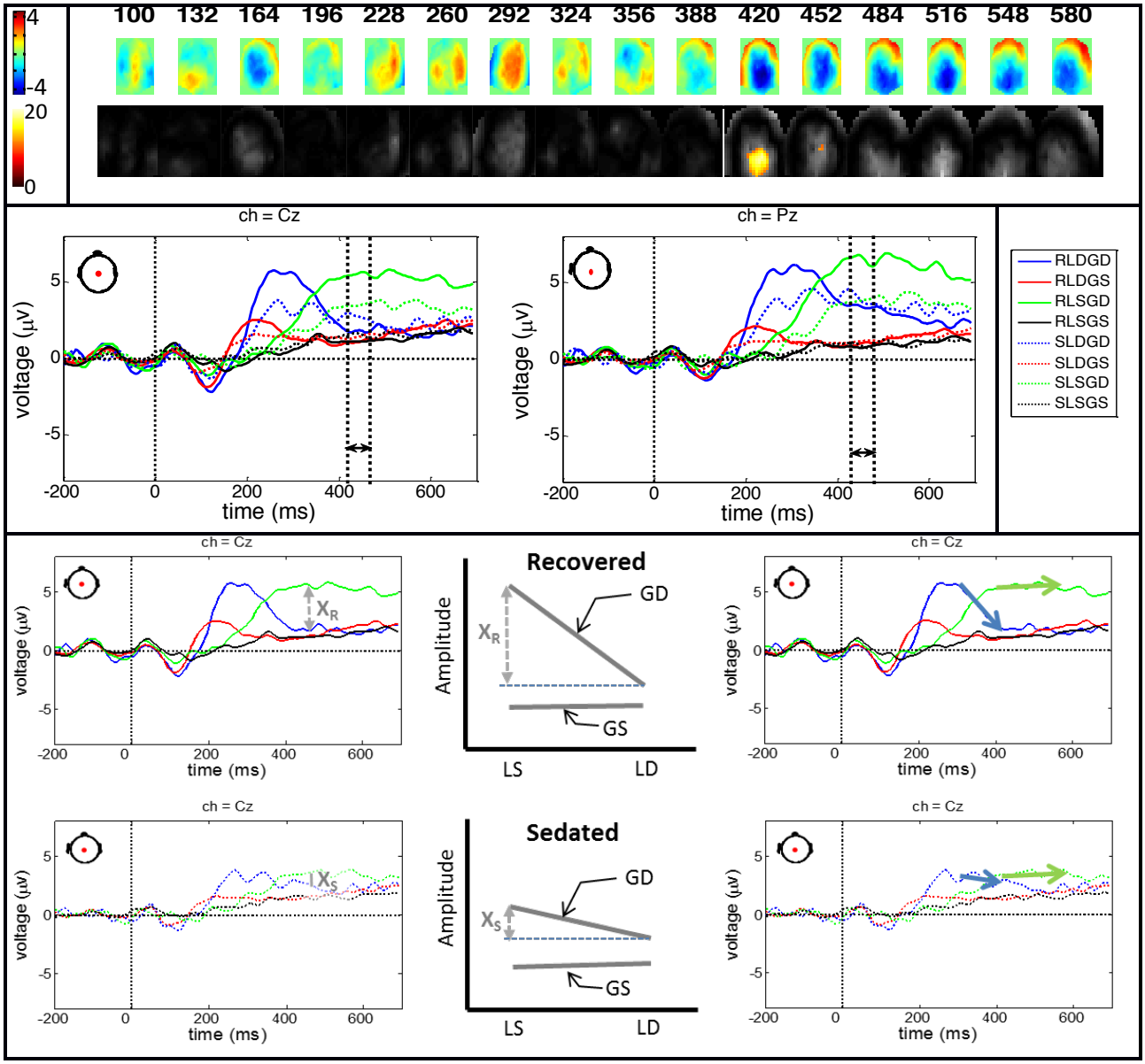


Figure 6. T-values, and thresholded F-values (output of the Random Field Theory (RFT)) for the 3 way interaction are presented in the top panel, the first and second rows, respectively. Sequences of scalp maps through time are shown, with front of head at the top of each scalp. Side channels are removed, cluster forming threshold is set at 0.001, and Family-wise Error (FWE) rate correction at 0.05. The corresponding ERPs for the 3-way interaction are depicted in the 2nd panel down. The double direction horizontal arrows show the time regions where the 3 way interaction is significant. The onset of the fifth tone is at zero, the time point from where the ERP responses are analysed

to find the effect. The window from -200 to 0ms is the baseline region. The bottom panel indicates the key differences in ERPs that drive the three-way interaction. The left side of the panel shows the 3-way interaction, broken into its two consistent 2-way interactions, i.e. ERPs for LxG when Recovered and LxG when Sedated (both at Cz). These same ERPs are repeated on the right side of the panel, with different annotation, i.e. arrows indicating the sharpness of offset. Interaction plots from the two constituent interactions are shown in the middle of the panel. Since RLDGS and RLSGS are effectively the same, and SLDGS and SLSGS are effectively the same, the difference in the (LxG) two-way interactions is driven by the difference between LSGD and LDGD conditions, which is bigger in the Recovered (X_R) than the Sedated (X_S) conditions. The two ERP plots on the right of the bottom panel suggest that it is the difference in the offset dynamics of the LDGD conditions that indirectly cause the three-way interaction, which happens later in time. Note, the sharp offset of recovered LDGD. LD (local deviant), LS (local standard), GD (global deviant), GS (global standard), R (recovered), S (sedated).

We can also see that the recovered global deviants are generally higher in amplitude than sedated global deviants; however, the key landmark that drives the change in the LxG between Recovered and Sedated is that the RLDGD P3 response is short-lived and sharper, appearing early and ending early. The right most two ERP plots of the bottom panel of Figure 6 show the change in latency difference between LDGD, and LSGD most clearly. The RLSGD condition rises to a stable amplitude around 400ms, see green arrow, a point by which RLDGD has effectively terminated at Cz (see blue arrow). Although the SLDGD also starts earlier than SLSGD, they both continue at similar amplitudes from 400ms onwards, see blue and green arrows (this

is particularly clear at Pz, where there is little evidence of termination of LDGD at all; see right ERP plot with eight conditions in middle panel). It is this change in the termination profile of LDGD when sedated that drives significance of the three way interaction. Further consideration of this 3-way interaction is given in the Supplementary Material, see section S2, “Three-way Interaction”.

Overall, it seems that sedation firstly, reduces the amplitude of the response to deviance, while secondly, decelerating the accelerated prediction error. That is, as evident in the ERPs in figure 6 for the 3-way interaction, the global deviant response seems to be accelerated by the presence of local deviance, i.e. RLDGD is earlier than RLSGD, and SLDGD is earlier than SLSGD. Sedation though seems to diminish this acceleration. Most noticeably, the P3 finishes more sluggishly in SLDGD than it does in RLDGD (see blue arrows in the ERP plots on the right of the bottom panel of figure 6 – top plot recovered and bottom plot sedated). As a result of this, RLDGD is substantially different from RLSGD, while SLDGD is not that different from SLSGD, when the three-way interaction is significant in the late window, at, for example, Cz and Pz; see “Three Way Interaction” subsection in Supplementary Material for further detailed justification of why this pattern generates a 3-way interaction.

3.2 Single-trial Analyses

The ERP analysis just reported suggests that responses when sedated are generally reduced in amplitude. As discussed in the Introduction and “Single Trial Analysis” subsection of the Materials and Methods, this reduction in amplitude at the ERP level could be due to either decreased single trial amplitude or that some trials are completely missed when sedated (what we are calling, the drifting tonic alertness

hypothesis). Therefore, responses may be completely missing on some trials, which would cause a reduction in the average effect (i.e. in an ERP analysis). To explore this, our single trial analysis tested the standard deviation of the distribution of response amplitudes when participants were sedated compared to when they were recovered.

3.2.1 Average amplitude analysis: Following the strategies discussed in the "Single Trial Analysis" subsection of the Materials and Methods section for window placement, we apply a window from 200ms to the end of the time region across replications in our first approach (Same window). A second approach (Varying Windows), with different windows across conditions, is also applied across replications; i.e. a window from 200ms to 400ms for LD conditions (i.e. RLDGD, SLDGD, RLDGS, SLDGS), and 250ms to the end for LS conditions (i.e. RLSGD, SLSGD, RLSGS, SLSGS). This approach was taken to place windows most appropriately for the presentation of the P3 in different conditions. The full factorial analysis (c.f. Supplementary Material, tables S.T2 to S.T5) gave no significant effect in relevant conditions involving the Sedation and Global factors.

Summary Table 1 for the sedation-global interaction presents the results based on the two window placement strategies at channel Cz and Pz. This provides no evidence to reject the null hypothesis that the reduced P3 amplitude when sedated is as consistently present at the single trial level as the higher P3 amplitude is when recovered. This however does not enable us to infer that participants present consistent behaviour, i.e. to affirm the null. Crucially, Bayes factors (BFs) can be used as a group level test to provide evidence in favour of the null; i.e. that there is no difference between conditions. Accordingly, we performed Bayesian

analysis for the two-way interaction (Dienes, 2014); i.e. the sedation-global interaction.

Strategy	Channel	df	F	Sig. (i.e. p value)	Partial Eta Squared
1	Cz	1, 17	.361	.556	.021
1	Pz	1,17	.244	.628	.014
2	Cz	1,17	2.992	.102	.150
2	Pz	1,17	.942	.345	.052

Table 1. Statistics for across-trial variation (standard deviation of amplitudes) between the global deviant and global standard modulated by sedation (i.e. sedation-global interaction). Replications windowed over 200ms to end of the time region in the first strategy (1), i.e. Same window, and windowed over 200ms to 400ms for LD conditions and 250ms to end for LS conditions in the second strategy (2), i.e. Varying windows.

3.2.2 Bayesian analysis: We first need to determine the range that the priors used in this analysis should take. As explained in the "Single Trial Analysis" subsection of the Materials and Methods section for prior estimation, we first calculate the limits of the 95% confidence interval for all conditions (i.e. RLDGD, RLDGS, RLSGD, RLSGS, SLDGD, SLDGS, SLSGD, SLSGS), to find the limits of the uniform prior. Using the t-distribution, Supplementary Material (tables S.T6 to S.T9) provides details of confidence intervals and average effects for each condition at channels Cz, and Pz, based on the two different window placement strategies. We give a worked example for this analysis in the Supplementary Material, section S5 "Confidence Intervals, Correction Factor and an Example for Bayesian Analysis". Details of the results of the Bayesian analysis are reported in Table 2. The ratio of the probability of the observed data under the alternative hypothesis to the probability of the

observed data under the null, and the resulting BF are presented in the last two columns of Table 2. The BFs, which are all below 0.3333, provides evidence in favour of the null in all four cases for the variation of single-trial average amplitudes. This suggests that the P3 variation across trials is as consistent when participants are sedated as when recovered.

Strategy/ Channel	F-value	t-value	Mean difference	SD-corrected	Prior limits	Hypothesis/ Null	B
1/Cz	.361	0.600	0.1518	0.2705	[-4 4]	0.1250/1.2600	0.0992
1/Pz	.244	0.493	0.1064	0.2307	[-3.5 3.5]	0.1429/1.5548	0.0919
2/Cz	2.992	1.729	0.4361	0.2697	[-4.5 4.5]	0.1111/0.4002	0.2776
2/Pz	.942	0.970	0.2142	0.2361	[-3.5 3.5]	0.1429/1.1197	0.1276

Table 2. Calculation of BFs with corrected standard error for the variation of single-trial average amplitudes using Dienes's calculator (2014). Standard error is corrected by the factor 1+20/(df*df). The first column presents the window placement strategies at different channels: Same window (1); and Varying windows (2). The second column presents the F-values. The third column (t-value) is the square root of the F-value. The forth column (mean difference) shows the mean of the interaction. The fifth column shows the corrected standard error. The sixth column shows the prior distribution limits. The seventh column (Hypothesis/Null) shows the BF ratio. The eighth column (B) is the BF, i.e. the result of Hypothesis/Null.

4. Discussion

4.1 Findings

We have presented a first full multi-factor ANOVA analysis of the local-global experiment using an ERP cluster-extent statistic. This has enabled us to assess the effect of sedation on a task that elicits a hierarchy of prediction errors – one at an early level (the local effect) and another that is later (the global effect).

4.1.1 Local effect: As seen in Figure 4, local effect panel (L), we were able to replicate the two basic characteristics of the local effect. 1) We found the previously observed (Bekinschtein et al., 2009) enhanced early (around 130ms) negativity to local deviant, which was largest at frontal electrodes. This corresponds to the Mismatch Negativity (Bekinschtein et al., 2009; Näätänen, Paavilainen, Rinne & Alho, 2007). 2) As previously (Bekinschtein et al., 2009; Chennu et al., 2013), we observed a (more central) positive rebound (beginning around 196ms) to the local effect. This positive rebound has been interpreted as a P3a (Naccache, King, Sitt, Engemann, & Karoui, 2015) [4].

4.1.2 Global effect: We also replicated the standard global effect (Bekinschtein et al., 2009), with a broad sustained posterior positive pattern; see Figure 4, global effect panel (G). This has the characteristic topography (midline, broad, centred close to Pz) and time window (reaching its peak around 400ms post fifth stimulus, c.f. Figure 4, global effect panel) of a P3b (Bekinschtein et al., 2009).

4.1.3 Local x global: A key contribution is our reporting and characterisation of an interaction between the local and global effects. Bekinschtein et al investigated, but

4 The positive (P3a) rebound was followed (around 420ms) by a small amplitude, but extended, negative effect at frontal sites, which lasted out towards the end of the analysed segment; see Figure 4, local effect panel (L). This reversal back to a negative effect is a new finding.

did not identify an interaction between these early and late effects (Bekinschtein et al., 2009). In contrast, we identified a strong two phase pattern, which was largest centrally; see Figure 5, LxG panel. This seems to be driven by an accelerated (and more temporally curtailed) response to global deviance, when paired with local deviance, while global standard conditions do not differ. That is, it seems that double (local and global) surprise elicits a swift and sharp registration of a prediction error passing up the auditory pathway[5].

A point to consider is that the identification of an interaction does not, in and of itself, provide a line of causation. Indeed, one might argue that our interaction is driven by global deviance accelerating the (P3a) positive rebound of the local deviance response. However, this seems a less compelling explanation, since the local effect is naturally interpreted as starting before the global effect, as reflected by the MMN

5 A local-global interaction was also observed by Wacongne et al. (2011). However, they reported it in an early time window: 134-190ms post fifth tone onset. This is described in the second paragraph of their Results section, where the difference between local deviant traces in lower left panels of their Figure 2A and 2B (purple trace in 2A and green in 2B) is large compared to the difference between local standards (green trace in 2A and purple in 2B). This is a time window in which we do not observe a significant interaction; ours begins around 260ms post fifth tone onset; see our Figure 5, the top panel, marked LxG. Thus, it would seem that the interactions we and Wacongne et al have observed are quite different.

A potential explanation of the Wacongne et al interaction is that it reflects the decay rate of auditory cortex. That is, the local deviant tone (the Y in XXXXY blocks, i.e. LDGS) recurs with a short enough delay between repetitions (one quintuple length plus the inter-quintuple interval) that the representation of Y builds up through such a block, causing it to become expected and thus predicted. This prediction is repeatedly confirmed, yielding a small PE to the local deviant. In contrast, a local deviant in an XXXXX block (i.e. LDGD) is not predicted, since it fully decays between presentations, generating a large local PE (such sensitivity of the MMN to the inter-quintuple interval is clear from the absence of an MMN in (Sussman, Ritter, & Vaughan, 1998) and presence of a small MMN in LDGS blocks in Bekinschtein et al. (2009).

This explanation of the Wacongne et al interaction is not particularly informative with regard to the global effect, since it arises as a function of the stimulus presentation sequence, rather than engagement of the later "global" processing stage, as is the essence of the LxG interaction we are reporting here. A related point to this is made in King et al. (2013), c.f. P733, beginning of section "Generalization across contexts isolates rule-specific effects".

Additionally, although not formulated in the terminology of an interaction, the finding in King et al. (2013) that classification does not generalize well across global effect contexts (e.g. from distinguishing GDLS from GSLS to distinguishing GDLD from GSLD) is consistent with our identification of a LxG interaction.

being so early an evoked response [6]. Furthermore, if one ascribed the accelerated positivity seen in LDGD to the local effect, one would be left without any global component at all in that condition (i.e. no EEG response reflecting the global irregularity), which would seem unlikely. It is for these reasons that we prefer the explanation that local deviance accelerates the response to global deviance [7].

An implication of identifying this interaction is that a clean division of the two effects, such that only the later (global) effect is a neural correlate of awareness, is less likely. The very interdependence of the two effects stands against them being, in an absolute sense, functionally and/or phenomenologically completely isolated from one another, i.e., local pre-conscious, global conscious.

4.1.4 Sedation: In general, responses are larger and more extreme when participants are recovered. That is, it is almost always the case that whenever ERPs are negative, recovered is smaller (i.e. more negative) and whenever positive, recovered is bigger (i.e. more positive); e.g. see ERP Figure 4, lower part of bottom panel, for the main effect of sedation. A particularly strong demonstration is in the N1, which, as previously identified (Yppärilä, Nunes, Korhonen, Partanen, & Ruokonen, 2004), is substantially reduced when sedated. Thus, and indeed unsurprisingly, sedation seems to broadly dampen the evoked response.

4.1.5 Sedation x local: Consistent with the local effect not being fully isolated from awareness, we observe an interaction between sedation and local; see Figure 5, SxL panel. This interaction occurs in the positive rebound of the local effect, which has a

6 This said, the local effect we observe does last on into a third phase, around 420ms at frontal sites (c.f. Figure 4, top panel), which continues out towards the end of the EEG segment. This later pattern is though relatively speaking, weaker, which, although extended in time, is low in amplitude.

7 It would, though, certainly be interesting to attempt to decompose the accelerated P3 in the LDGD condition with say ICA or PCA to see if there is any sense to which two subcomponents are present, with one associated with the local and the other the global effect.

Cz distribution, around 228 ms. Specifically, sedation has very little effect on the local standard response in this time-space region, but a differential effect on local deviance, inducing a selective dampening of the brain's second phase error response for local prediction.

This finding seems somewhat consistent with past human data (Heinke, Fiebach, et al., 2004; Heinke, Kenntner, et al., 2004) and those acquired from non-human primates (Uhrig et al., 2016), where propofol does not attenuate (or may even enhance) the primary perceptual activation associated with an auditory stimulus, but clearly attenuates the incremental local activation produced by locally deviant stimuli. In our data, this argument would particularly be focussed on the incremental local activation associated with the positive rebound (which may engage inferior frontal areas) of the local effect. This is important, as the summed local ERP response now represents the differential neural input that modulates processing in subsequent layers, and a reduction in such input could attenuate facilitatory modulation of a subsequent global response; see our discussion to follow under "Candidate Explanation" to Figure 7.

4.1.6 Sedation x global: Sedation reduces the global effect. Once again anaesthesia has little, if any, effect on the standard conditions, see ERPs in Figure 5, SxG panel. In contrast, global deviant recovered is higher amplitude and slightly earlier than global deviant sedated. Thus, sedation dampens (and slightly delays) the global prediction error response,

4.1.7 Local x global x sedation: Importantly, sedation with propofol modulates the interaction between the local and global effects; see Figure 6. This arises late in time, around 420ms post eliciting stimulus, and with a midline centro-posterior

pattern**.** Essentially, this interaction is driven by a reduction in the extent to which local deviance advances and curtails the evoked response to global deviance. That is, coincidence of a local prediction error temporally facilitates the response associated with the global prediction error (our local x global interaction shows this); however, the extent of this facilitation reduces when sedated. In contrast, during the space-time region of three-way significance, there is no modulation of the response to globally predictable auditory stimulation (i.e. global standard) by either local predictability or sedation. This combination creates the three-way significance. This deceleration of the accelerated prediction error is particularly evident in the ERPs of Figure 6.

4.1.8 Single trial analysis: Finally, out single trial analyses stand against the possibility that the amplitude reductions we observe when sedated, and especially those for the global effect, are caused by drifting tonic alertness. Using a Bayesian analysis, we were able to find evidence in favour of the null hypothesis that the amplitude of the global effect is as consistent across replications when sedated as when recovered.

4.2 Prediction Error and Free Energy

The key contribution of our work for predictive coding (Friston, 2010) is a demonstration that confounding of expectation can impact the *temporal responsiveness* of units. That is, an increase in prediction error does not just elevate the amplitude of the evoked response, it can also accelerate it.

A deep link has been proposed between PEs and evoked responses. Although other factors, such as, relevance to goals, task set and affective salience (Bowman, Filetti, Wyble, & Olivers, 2013) seem to contribute, the unexpected is undoubtedly a major

driver of evoked responses (Bridwell, Kiehl, Pearlson, & Calhoun, 2014; Kutas & Federmeier, 2011; Näätänen, Pakarinen, Rinne, & Takegata, 2004; Todd, Michie, Schall, Ward, & Catts, 2012).

From this perspective, as previously reported (Bekinschtein et al., 2009), we observed large amplitude increases in evoked responses with PEs, e.g. for the local and also for the global main effects, c.f. Figure 4, top and middle panels. Most notably, though, the combination of multiple levels of PE, i.e. the local-global interaction, is associated with a dramatic acceleration of the evoked response (i.e. for LDGD compared to LSGD), with much less evidence of an amplitude enhancement, c.f. Figure 5, lower part of LxG panel. This is what we call the *double surprise acceleration effect*.

This would seem to suggest some sort of *urgency* associated with the predictive coding framework. That is, a multilevel confounding of expectation is speeded, presumably because it is more noteworthy to the system, has a higher priority and thus needs to be registered more swiftly. This would imply that surprise is, in a sense, not locally encapsulated; that is, a stimulus that is locally unexpected increases the system's efficiency in detecting other levels of prediction error.

As an illustration, consider the architecture in panel C of Figure 1. Although there is an output-input relationship between local and global circuits, this is assumed to be a simple synaptic transmission, with multiplication by, what is, importantly, a *fixed* weight. Our contention is that such a simple system with fixed weights is insufficient to explain the observed LxG interaction. An additional modulatory pattern of influence from local to global is required. We elaborate on this shortly; c.f. subsection "Candidate Explanation".

4.3 Sedation

Our findings suggest the following, 1) propofol broadly reduces PEs, as indexed by evoked EEG responses to confounding of expectation. These reductions are observed early in the processing pathway, when integrating over fine temporal contexts (local effect) and later, when integrating over coarser contexts (global effect). 2) Significantly, the acceleration of the global evoked response by local deviance is partially counteracted by sedation; that is, sedation decelerates the acceleration generated by double surprise. 3) These modulatory effects of propofol seem *not to be* driven by substantial waxing and waning of tonic alertness; that is, the modulations consistently obtain at the single trial level.

It is important to recognise that propofol has several mechanisms of action (Trapani, Altomare, Liso, Sanna, & Biggio, 2000; Vanlersberghe & Camu, 2008), e.g. potentiation of GABAA receptor activity, thereby slowing the channel closing time (Trapani et al., 1998), action through the endocannabinoid system (Fowler, 2004), sodium channel blocking (Haeseler et al., 2008) and impact on gamma frequencies in EEG (Lee, Mashour, Kim, Noh, & Choi, 2009). Consequently, we are unable to definitively identify the neurophysiological action that underlies the effects we are observing. Nonetheless, propofol's sedative effect is clear, indicating a broad impact on the state of consciousness.

Additionally, the data analysed here were acquired as part of a larger experimental protocol, which precluded acquisition of responses to the relevant stimuli in the awake state – data were only obtained while the participant was sedated (Sedation), and then when they had behaviourally recovered from sedation (Recovered). The comparison to an awake condition is not critical to the hypothesis we wished to test,

since all that the analysis demanded was measurement of neural responses in conditions with a difference in level of consciousness – which was provided by the Sedation and Recovered conditions.

Additionally, participants were instructed to attend to global deviance, and they could be expected to at the least be attempting to maintain such an attentional state, even while the anaesthetic was having its fullest effect. Consequently, while our findings inform the question of how *reduced* awareness impacts evoked responses, they certainly do not conclusively answer the question of how such responses manifest in the absolute absence of awareness. Accordingly, our findings do not definitively contradict the position that the global effect can only be observed in conscious attentive participants; see (Naccache et al., 2015) for a clear discussion of this point. However, from this position, our finding that sedation has a wide effect, including on both local and global, may be considered noteworthy. See Chennu & Bekinschtein (2012) for a relevant discussion of these issues.

A possible explanation for our finding of a Local x Global x Sedation interaction is that propofol alters the ability of critical higher centres or systems to undergo modulation by neural inputs from lower processing layers generated by a locally deviant stimulus. The generation of a response to a globally deviant stimulus is believed to engage a large network of brain regions, with substantial contributions from the frontal lobes (as shown by both fMRI and source localisation data; (Bekinschtein et al., 2009)). It is well recognised that propofol more selectively affects frontal regions than other intravenous anaesthetics (Veselis et al., 2004), with reduced fMRI activation (Heinke & Fiebach et al., 2004) in response to simple stimuli and globally deviant stimuli (Uhrig et al., 2016).

More broadly, propofol may reduce connectivity within regions responsible for neural responses to locally or globally deviant stimuli, or through communication between these regions. The available data (Guldenmund et al., 2016; Ku, Lee, Noh, Jun, & Mashour, 2011; Lee, Kim, et al., 2009; Liu et al., 2012) strongly support such modulation of connectivity by anaesthetic agents, both in humans and non-human primates. Interestingly, the suggestion from some studies (Guldenmund et al., 2016; Lee, Kim, et al., 2009) is that feed-forward and feed-back connectivity between frontal and parietal regions is differentially affected by propofol sedation and anaesthesia, with feedback connectivity attenuated at lower levels of anaesthesia. Indeed, the attenuation of frontoparietal long range connectivities with propofol sedation, which occurs at low levels of sedation, may be responsible for early cognitive effects of the drug, and precedes the attenuation of thalamocortical connectivity that more specifically underpins loss of responsiveness (Barttfeld et al., 2015).

The literature then leaves two particularly prominent potential explanations for our LxGxS interaction. That is, the deceleration of the P3 in the LDGD condition when sedated, which drives our three-way interaction, may arise due to A) a loss of frontal responsiveness and/ or B) a decline in the integrated action of the fronto-parietal network (which supports the global effect) due to connectivity changes. In either case, the responsiveness of areas supporting the putative GW would be impaired.

Broadly speaking, then, our data are concordant with past studies, which show that the early response to locally deviant stimuli and the later response to globally deviant stimuli are both attenuated by propofol sedation. However, we additionally show that the facilitation by responses to locally deviant stimuli of later responses to globally deviant stimuli, associated with conscious access, are attenuated by propofol.

4.4 Two Modes for Detecting Unexpected Sounds – Distinct or Indistinct?

It has been suggested that the local-global task reveals two distinct processing modes – the first with the character of a prediction system, as per Friston's predictive coding theory, and the second with the character of Dehaene et al's Global Workspace (King et al., 2014). Supported by impressive MEEG decoding analyses, King et al have argued that the first of these is engaged by the local and the second the global effect. This position has important theoretical implications, since it is argued to expose the limits of predictive coding. Specifically, it suggests that a "meta-stable" sustained activation system is required to explain the temporally extended integration and storage associated with the working memory demands of the global effect. Such extended states are argued to be critical to conscious experience and to be beyond the basic predictive coding hierarchy.

We have investigated the distinctness of these two modes by looking for interactions between the local and global effects, while our sedation manipulation enables us to explore how consciousness modulates these two modes, individually and in their interaction. Accordingly, our findings cast doubt on the absolute distinctness of the two modes. While the necessity for a more sustained representation scheme for realising the global effect is clear, modulatory isolation of this scheme from the local system seems less certain. We observed that local deviance had a strong modulatory effect on the response to global deviance, dramatically accelerating the time-course of the resulting P3; see Figure 5, LxG panel.

In addition, and perhaps most significantly, our results suggest that the stability of the global state is modulated by the coincidence of local regularity. That is, it seems that the sustaining of the global state is increased by confirmation of local

predictions, or equivalently, the global state is curtailed in the presence of a local prediction error, see ERP Figure 5, LxG panel.

As a result, one might argue that the interdependence we observe between hierarchical levels is suggestive of the multi-directional exchanges associated with the predictive coding architecture. Although, exactly how consistent the cross-level dependences we observe are with the specifics of predictive coding awaits dynamic causal modelling (Stephan et al., 2008) on this data set.

Finally, unique association of conscious experience with the global state is challenged by our findings. Firstly, sedation was found to weaken the local as well as the global effect; c.f. the SxL panel in Figure 5, second row. Secondly, and most interestingly, sedation effectively reduced the extent to which the local effect multiplicatively modulated the global effect; c.f. Figure 6. This suggests firstly that the local effect directly impacts the “control” dynamics of the global circuit, and furthermore, that that influence seems to be specifically predicated upon (or at the least heightened by) conscious experience.

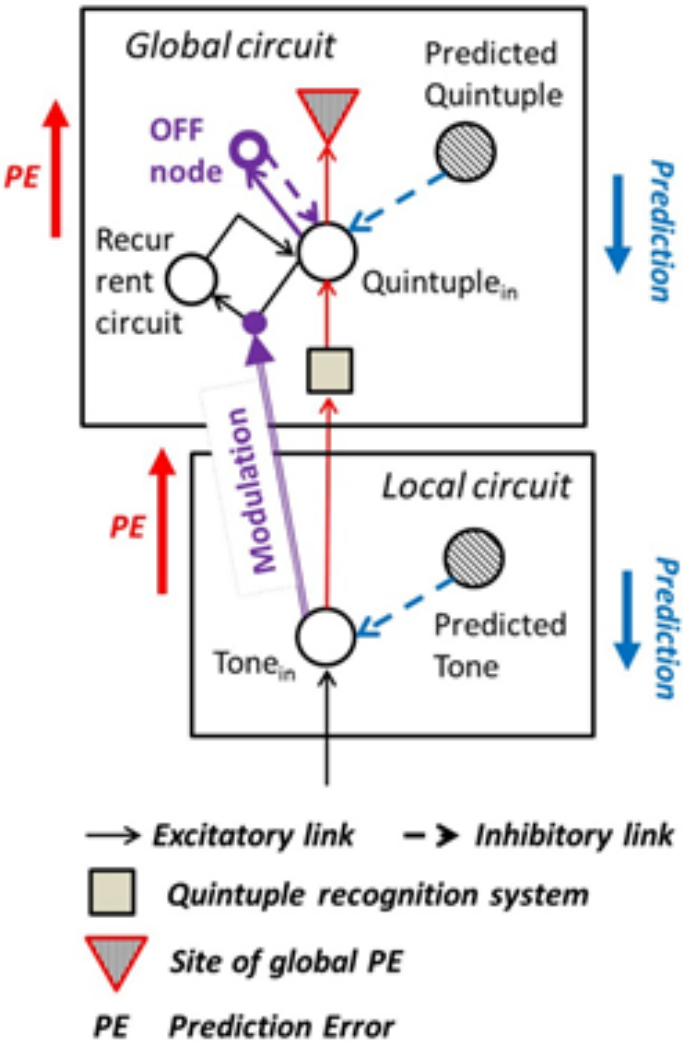


Figure 7. Sketch of a neural hierarchical prediction architecture that could realise the double surprise acceleration effect.

4.5 Candidate Explanation

Although it is certainly no more than a candidate explanation, it is valuable to consider how the key findings of this paper might be realised in a neural architecture. We do this by adding just two mechanisms to the network highlighted at the start of this paper, in Figure 1, panel C [8]. The resulting architecture is shown in Figure 7,

8 The architecture we propose here is in no way intended to serve as a complete computational model of the phenomena we are considering – it is no more than a sketch of a model. The impressive neural network

with the added mechanisms shown in purple. These are 1) a multiplicative modulation of the global recurrent circuit from the local level, and 2) an attractor termination mechanism realised with an OFF node. The key new findings of this paper are assumed to arise in the following ways.

4.5.1 Sedation: Sedation could generate an across-the-board reduction in excitatory influences, broadly decreasing the level of activation reached by excitatory units. This may be particularly pronounced during the positive rebound of the local effect (c.f. our Sedation x Local pattern) and relatively early in the P3 (c.f. our Sedation x Global pattern). (See justification in subsection "Sedation" of this Discussion).

4.5.2 Local x Global interaction: The double surprise acceleration effect would arise for two reasons. Firstly, a global deviant that is also locally deviant would not be opposed by an active quintuple prediction unit (in contrast to a global standard). Secondly, it would benefit from the recurrent excitation circuit being strong, which would accelerate the P3, driving that assembly of units more swiftly to an attractor state. This circuit strength would be ensured by the modulatory effect of a local deviant on the global level recurrent circuit (via the modulation projection).

Thus, registration of local deviance might have a nonlinear modulatory effect on the assembly of units coding the global deviance response. Multiplicative modulation of units on other parts of a network can be found in a number of modelling frameworks. For example, some more expressive classes of Dynamic Causal Modelling incorporate nonlinear projections from units onto link weights (Stephan et al., 2008). These could dynamically enhance coupling, making a recurrently interconnected assembly more temporally responsive.

---

simulation of auditory deviance of Wacongne, Changeux, & Dehaene (2012) offers a much more complete computational theory, although it is a theory specific to the local level of the local-global task. In contrast, our main focus is the global effect and the interfacing of it with the local effect.

The flip side of this is that if a stimulus is locally expected, global unexpectedness (i.e. the LSGD condition) does not get registered so quickly. The network in Figure 7 would generate this pattern, since, in that case, the $Tone_{in}$ unit would be suppressed by a strong Tone prediction unit, weakening the excitatory modulation of the global recurrent circuit. As a result, the global response would be slowed, as per the LSGD condition.

An intriguing aspect of our ERPs is that, although double surprise advances the global prediction error in time, it also seems to terminate the response quickly. A comparison of the ERPs for LDGD and LSGD in the lower row of top (LxG) panel in Figure 5 shows this, with the LDGD deflection terminating sharply around 350ms, while the LSGD deflection starts later, but continues out to the end of the analysed segment. A means to terminate a strong activation is to include an "off" loop (Bowman, Schlaghecken, & Eimer, 2006; Bowman & Wyble, 2007), such as we have incorporated in our global circuit in Figure 7. The OFF node would integrate the activation of the $Quintuple_{in}$ unit through time, crossing threshold when that accumulation of activation itself crossed a particular level. Accordingly, activation in the recurrent circuit would be terminated sharply once it had activated the OFF node sufficiently. Indeed, the sharp offset of LDGD activation may be suggestive of exactly such an OFF node "cutting-in".

4.5.3 Local x Global x Sedation: deceleration of the double surprise acceleration effect through sedation, suggests that Propofol reduces the capacity for local deviance to modulate global temporal responsiveness. A simple way to obtain this in the candidate architecture would be to weaken the excitatory recurrence circuit. This could be done by reducing the responsiveness to (or indeed of) the modulatory projection from the local level to the circuit, decelerating the evoked response to a

global deviant, when paired with a local deviant. The mechanisms highlighted in subsection “Sedation” of this “Discussion” concerning changes in the responsiveness of the fronto-parietal network, could clearly play a role here.

Additionally, our findings could be interpreted as suggesting that reduced awareness (as per the sedation condition) impairs the shutting off of the global PE response, with this being evident in the double surprise (LDGD) condition.

In conclusion, then, it does seem that (hierarchically) layering surprise upon surprise accelerates registration of the ensuing prediction error, and that sedation decelerates that registration.

Funding

This work was supported by grants from the James S. McDonnell Foundation, the Wellcome Trust [WT093811MA to TAB], the British Oxygen Professorship from the Royal College of Anaesthetists [to DKM], the UK Engineering and Physical Sciences Research Council [EP/P033199/1 to SC] and the Evelyn Trust [15/07 to SC]. The research was also supported by the NIHR Brain Injury Healthcare Technology Co-operative based at Cambridge University Hospitals NHS Foundation Trust and the University of Cambridge. The views expressed are those of the authors and not necessarily those of the UK National Health Service, the NIHR or the UK Department of Health. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Acknowledgments

We would like to thank the SPM research group at the Wellcome Trust Centre, UCL, especially Vladimir Litvak for considerable advice. We would also like to thank Karl Friston and Sebastian Marti for illuminating discussions, which informed the theoretical position taken in this paper. Finally, we also thank Brad Wyble for advice on filtering, as well as Elkan Akyurek and an anonymous reviewer for very thoughtful reviewer suggestions on the paper.

Address correspondence to Howard Bowman, School of Psychology, University of Birmingham, Edgbaston, Birmingham B15 2TT.

5. References

Bareham, C. A., Manly, T., Pustovaya, O. V, Scott, S. K., & Bekinschtein, T. A. (2014). Losing the left side of the world : Rightward shift in human spatial attention with sleep onset, 1–5. https://doi.org/10.1038/srep05092

Barttfeld, P., Bekinschtein, T. A., Salles, A., Stamatakis, E. A., Adapa, R., Menon, D. K., & Sigman, M. (2015). Factoring the brain signatures of anesthesia concentration and level of arousal across individuals. *NeuroImage: Clinical*, *9*, 385–391. https://doi.org/10.1016/j.nicl.2015.08.013

Bekinschtein, T. A., Dehaene, S., Rohaut, B., Tadel, F., Cohen, L., & Naccache, L. (2009). Neural signature of the conscious processing of auditory regularities. *Proceedings of the National Academy of Sciences*, *106*(5), 1672–1677. https://doi.org/10.1073/pnas.0809667106

Bowman, H., Filetti, M., Wyble, B., & Olivers, C. (2013). Attention is more than prediction precision. *Behavioral and Brain Sciences*, *36*(03), 206–208. https://doi.org/10.1017/S0140525X12002324

Bowman, H., Schlaghecken, F., & Eimer, M. (2006). A neural network model of inhibitory processes in subliminal priming. *Visual Cognition*, *13*(4), 401–480. https://doi.org/10.1080/13506280444000823

Bowman, H., & Wyble, B. (2007). The simultaneous type, serial token model of temporal attention and working memory. *Psychological Review*, *114*(1), 38–70. https://doi.org/10.1037/0033-295X.114.1.38

Bridwell, D. A., Kiehl, K. A., Pearlson, G. D., & Calhoun, V. D. (2014). Patients with schizophrenia demonstrate reduced cortical sensitivity to auditory oddball regularities. *Schizophrenia Research*, *158*(1), 189–194. https://doi.org/10.1016/j.schres.2014.06.037

Brooks, J., Zoumpoulaki, A., & Bowman, H. (2016). Data-driven region-of-interest selection without inflating Type I error rate. *Psychophysiology*.

Chennu, S., & Bekinschtein, T. A. (2012). Arousal Modulates Auditory Attention and Awareness: Insights from Sleep, Sedation, and Disorders of Consciousness. *Frontiers in Psychology*, *3*, 65. https://doi.org/10.3389/fpsyg.2012.00065

Chennu, S., Noreika, V., Gueorguiev, D., Blenkmann, A., Kochen, S., Ibáñez, A., … Bekinschtein, T. A. (2013). Expectation and Attention in Hierarchical Auditory Prediction. *Journal of Neuroscience*, *33*(27).

Dayan, P., & Abbott, L. F. (2001). Theoretical neuroscience, vol. 806. Cambridge, MA: MIT Press.

Dehaene, S., Kerszberg, M., & Changeux, J. P. (1998). A neuronal model of a global workspace in effortful cognitive tasks. *Proceedings of the National Academy of Sciences of the United States of America*, *95*(24), 14529–34. https://doi.org/10.1073/PNAS.95.24.14529

Dienes, Z. (2014). Using Bayes to get the most out of non-significant results. *Frontiers in Psychology*, *5*, 781. https://doi.org/10.3389/fpsyg.2014.00781

Eklund, A., Nichols, T. E., & Knutsson, H. (2016). Cluster failure: Why fMRI inferences for spatial extent have inflated false-positive rates. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(28), 7900–5. https://doi.org/10.1073/pnas.1602413113

Feldman, H., & Friston, K. (2010). Attention, uncertainty, and free-energy. *Frontiers in Human Neuroscience*.

Flandin, G., & Friston, K. J. (2016). Analysis of family-wise error rates in statistical parametric mapping using random field theory.

Fowler, C. J. (2004). Possible involvement of the endocannabinoid system in the actions of three clinically used drugs. *Trends in Pharmacological Sciences*. https://doi.org/10.1016/j.tips.2003.12.001

Friston, K. (2005). A theory of cortical responses. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, *360*(1456).

Friston, K. (2010). The free-energy principle: a unified brain theory? *Nature Reviews Neuroscience*, *11*(2), 127–138. https://doi.org/10.1038/nrn2787

Friston, K. J., Price, C. J., Fletcher, P., Moore, C., Frackowiak, R. S. J., & Dolan, R. J. (1996). The Trouble with Cognitive Subtraction. *NeuroImage*. https://doi.org/10.1006/nimg.1996.0033

Friston, K., & Kiebel, S. (2009). Predictive coding under the free-energy principle. *Philosophical Transactions Of*.

Guldenmund, P., Gantner, I. S., Baquero, K., Das, T., Demertzi, A., Boveroux, P., … Soddu, A. (2016). Propofol-Induced Frontal Cortex Disconnection: A Study of Resting-State Networks, Total Brain Connectivity, and Mean BOLD Signal Oscillation Frequencies. *Brain Connectivity*. https://doi.org/10.1089/brain.2015.0369

Haeseler, G., Karst, M., Foadi, N., Gudehus, S., Roeder, A., Hecker, H., … Leuwer, M. (2008). High-affinity blockade of voltage-operated skeletal muscle and neuronal sodium channels by halogenated propofol analogues. *British Journal of Pharmacology*. https://doi.org/10.1038/bjp.2008.255

Heinke, W., Fiebach, C. J., Schwarzbauer, C., Meyer, M., Olthoff, D., & Alter, K. (2004). Sequential effects of propofol on functional brain activation induced by auditory language processing: an event-related functional magnetic resonance imaging study. *British Journal of Anaesthesia*. https://doi.org/10.1093/bja/aeh133

Heinke, W., Kenntner, R., Gunter, T. C., Sammler, D., Olthoff, D., & Koelsch, S. (2004). Sequential effects of increasing propofol sedation on frontal and temporal cortices as indexed by auditory event-related potentials. *Anesthesiology*.

Kanai, R., Komura, Y., Shipp, S., & Friston, K. (2015). Cerebral hierarchies: predictive processing, precision and the pulvinar. *Philosophical Transactions of the Royal Society B: Biological Sciences*. https://doi.org/10.1098/rstb.2014.0169

Kilner, J. (2013). Bias in a common EEG and MEG statistical analysis and how to avoid it. *Clinical Neurophysiology*.

King, J. R., Faugeras, F., Gramfort, A., Schurger, A., El Karoui, I., Sitt, J. D., … Dehaene, S. (2013). Single-trial decoding of auditory novelty responses facilitates the detection of residual consciousness. *NeuroImage*. https://doi.org/10.1016/j.neuroimage.2013.07.013

King, J. R., Gramfort, A., Schurger, A., Naccache, L., & Dehaene, S. (2014). Two distinct dynamic modes subtend the detection of unexpected sounds. *PLoS ONE*. https://doi.org/10.1371/journal.pone.0085791

Kriegeskorte, N., Simmons, W. K., Bellgowan, P. S. F., & Baker, C. I. (2009). Circular analysis in systems neuroscience: the dangers of double dipping—Supplementary text. *Nature Neuroscience*. https://doi.org/10.1038/nn.2303

Ku, S., Lee, U., Noh, G., Jun, I., & Mashour, G. (2011). Preferential inhibition of frontal-to-parietal feedback connectivity is a neurophysiologic correlate of general anesthesia in

surgical patients. *PloS One*.

Kutas, M., & Federmeier, K. (2011). Thirty years and counting: finding meaning in the N400 component of the event-related brain potential (ERP). *Annual Review of Psychology*.

Lee, U., Kim, S., Noh, G. J., Choi, B. M., Hwang, E., & Mashour, G. A. (2009). The directionality and functional organization of frontoparietal connectivity during consciousness and anesthesia in humans. *Consciousness and Cognition*. https://doi.org/10.1016/j.concog.2009.04.004

Lee, U., Mashour, G. A., Kim, S., Noh, G. J., & Choi, B. M. (2009). Propofol induction reduces the capacity for neural information integration: Implications for the mechanism of consciousness and general anesthesia. *Consciousness and Cognition*. https://doi.org/10.1016/j.concog.2008.10.005

Liu, X., Lauer, K. K., Ward, B. D., Rao, S. M., Li, S. J., & Hudetz, A. G. (2012). Propofol disrupts functional interactions between sensory and high-order processing of auditory verbal memory. *Human Brain Mapping*. https://doi.org/10.1002/hbm.21385

Marsh, B., White, M., Morton, N., & Kenny, G. N. C. (1991). Pharmacokinetic model driven infusion of propofol in children. *British Journal of Anaesthesia*. https://doi.org/10.1093/bja/67.1.41

May, P., & Tiitinen, H. (2010). Mismatch negativity (MMN), the deviance-elicited auditory deflection, explained. *Psychophysiology*.

Näätänen, R., Gaillard, A., & Mäntysalo, S. (1978). Early selective-attention effect on evoked potential reinterpreted. *Acta Psychologica*.

Näätänen, R., Paavilainen, P., Rinne, T., & Alho, K. (2007). The mismatch negativity (MMN) in basic research of central auditory processing: a review. *Clinical Neurophysiology*.

Näätänen, R., Pakarinen, S., Rinne, T., & Takegata, R. (2004). The mismatch negativity (MMN): towards the optimal paradigm. *Clinical Neurophysiology*.

Naccache, L., King, J.-R., Sitt, J., Engemann, D., El Karoui, I., Rohaut, B., … Dehaene, S. (2015). Neural detection of complex sound sequences or of statistical regularities in the absence of consciousness? *Brain*, *138*(12), e395–e395. https://doi.org/10.1093/brain/awv190

Shirazibeheshti, A. (2015). The effect of sedation on conscious processing: Computational analysis of the EEG response to auditory irregularity, unpublished PhD thesis, University of Kent.

Stephan, K. E., Kasper, L., Harrison, L. M., Daunizeau, J., den Ouden, H. E. M., Breakspear, M., & Friston, K. J. (2008). Nonlinear dynamic causal models for fMRI. *NeuroImage*. https://doi.org/10.1016/j.neuroimage.2008.04.262

Sussman, E., Ritter, W., & Vaughan, H. (1998). Attention affects the organization of auditory input associated with the mismatch negativity system. *Brain Research*.

Todd, J., Michie, P. T., Schall, U., Ward, P. B., & Catts, S. V. (2012). Mismatch negativity (MMN) reduction in schizophrenia-Impaired prediction-error generation, estimation or salience? *International Journal of Psychophysiology*. https://doi.org/10.1016/j.ijpsycho.2011.10.003

Trapani, G., Altomare, C., Liso, G., Sanna, E., & Biggio, G. (2000). Propofol in anesthesia. Mechanism of action, structure-activity relationships, and drug delivery. *Current Medicinal Chemistry*. https://doi.org/10.2174/0929867003375335

Trapani, G., Latrofa, A., Franco, M., Altomare, C., Sanna, E., Usala, M., … Liso, G. (1998). Propofol analogues. Synthesis, relationships between structure and affinity at GABA(A) receptor in rat brain, and differential electrophysiological profile at recombinant human GABA(A) receptors. *Journal of Medicinal Chemistry*. https://doi.org/10.1021/jm970681h

Uhrig, L., Janssen, D., Dehaene, S., & Jarraya, B. (2016). Cerebral responses to local and global auditory novelty under general anesthesia. *NeuroImage*.

Vanlersberghe, C., & Camu, F. (2008). Propofol. In *Modern Anesthetics*. https://doi.org/10.1007/978-3-540-74806-9_11

Veselis, R. a, Feshchenko, V. a, Reinsel, R. a, Dnistrian, A. M., Beattie, B., & Akhurst, T. J. (2004). Thiopental and propofol affect different regions of the brain at similar pharmacologic effects. *Anesthesia and Analgesia*. https://doi.org/10.1213/01.ANE.0000131971.92180.DF

Wacongne, C., Changeux, J. P., & Dehaene, S. (2012). A Neuronal Model of Predictive Coding Accounting for the Mismatch Negativity. *Journal of Neuroscience*. https://doi.org/10.1523/JNEUROSCI.5003-11.2012

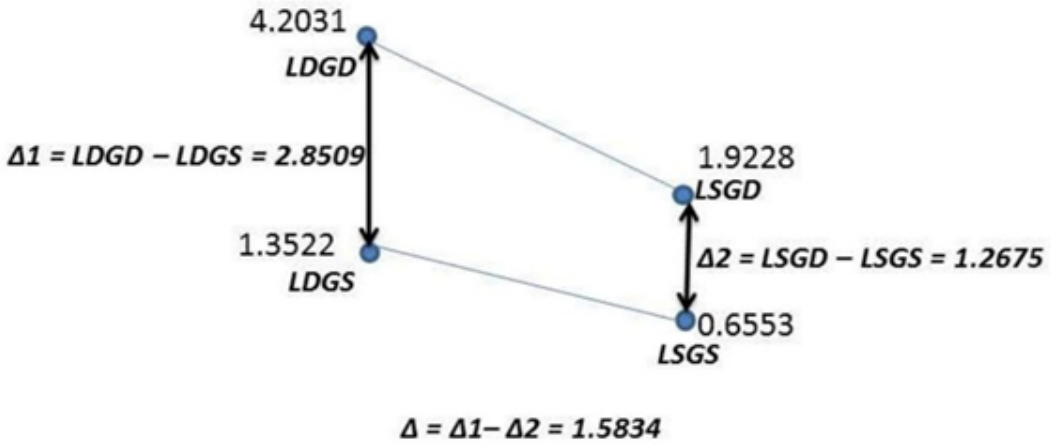
Wacongne, C., Labyt, E., van Wassenhove, V., Bekinschtein, T., Naccache, L., & Dehaene, S. (2011). Evidence for a hierarchy of predictions and prediction errors in human cortex. *Proceedings of the National Academy of Sciences*. https://doi.org/10.1073/pnas.1117807108

Woo, C., Krishnan, A., & Wager, T. (2014). Cluster-extent based thresholding in fMRI analyses: pitfalls and recommendations. *Neuroimage*.

Yppärilä, H., Nunes, S., Korhonen, I., Partanen, J., & Ruokonen, E. (2004). The effect of interruption to propofol sedation on auditory event-related potentials and electroencephalogram in intensive care patients. *Critical Care (London, England)*. https://doi.org/10.1186/cc2984

Supplementary Material

S1. Local-global Interaction. To explain the LxG interaction further, we calculated the means and differences that underlie the local-global interaction at one time-point (300) for channel Cz. In order for the local and global effects to be independent of one another, the difference between global deviant and global standard responses should be the same regardless of the level of the local condition (deviant vs. standard). Our findings are indicative of an interaction, with the relationship between the two factors being multiplicative, and not additive (figure S1).



LDGD = 4.2031, LDGS = 1.3522, LSGD = 1.9228, LSGS = 0.6553

Figure S1. The calculation of the local-global interaction at channel Cz, at 300ms, where the effect is significant. Δ1 = LDGD-LDGS, Δ2 = LSGD-LSGS, and interaction is Δ = Δ1 – Δ2 (the difference of differences). Δ=1.5834. Thus, the interaction (difference of differences), which is (LDGD-LDGS) - (LSGD-LSGS), is 1.5834 at this point. In and of itself, the size of this difference of differences does not guarantee significance, since the variance at this point will impact significance, as well as, the Random Field Theory (RFT), which takes into account cluster extent through time and space, but nonetheless the

pattern underlying this cluster is the one shown in figure 5 (main body of paper) top panel for LxG.

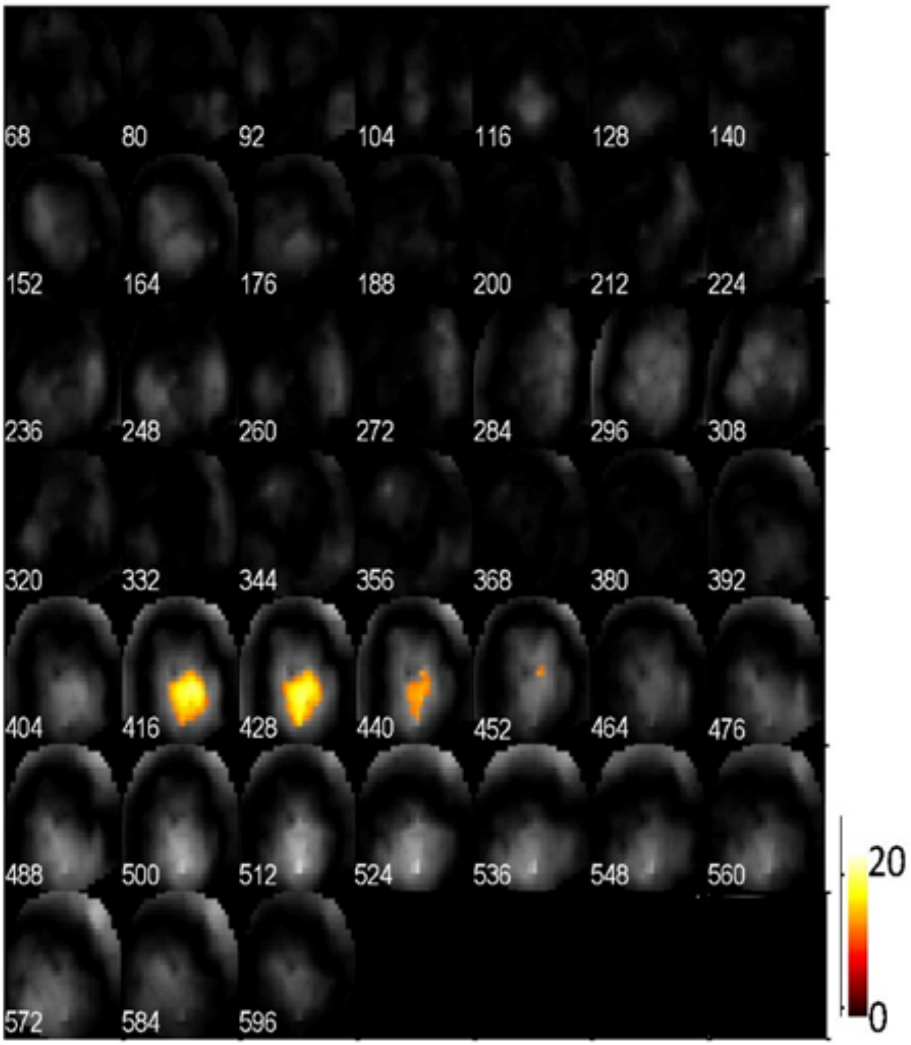


Figure S2. Thresholded F-maps for three way interaction with a fine resolution (12ms). The front of head is at the top of each scalp map. Side channels are removed, cluster forming threshold is set at 0.001, and Family-wise Error (FWE) rate correction at 0.05.

S2. Three-way Interaction. The three-way interaction can be viewed as the difference of two other interactions. The first is the local-global when recovered; i.e. σ1 = (RLDGD-RLDGS)-(RLSGD-RLSGS), and the second is the local-global when sedated; i.e. σ2 = (SLDGD-SLDGS)-(SLSGD-SLSGS). Then, the 3-way interaction is

simply the difference of these two differences; i.e. σ = σ1- σ2. Taking into consideration that the global standard changes very little between conditions within the activation period, and would largely cancel out of the calculation of σ, one may estimate the 3-way interaction as:

σ = (RLDGD-RLSGD)-(SLDGD-SLSGD) (2)

Clearly, the first part of σ, i.e. RLDGD-RLSGD, has a larger absolute amplitude than the second part, i.e. SLDGD-SLSGD. This is because, in the significant region, RLSGD is among the highest amplitudes, whereas RLDGD is among the lowest amplitudes of the four conditions. This is what drives the 3-way interaction.

To show its topographical form more fully, figure S2 presents the pattern of three way activation over time with a fine resolution.

Nature of 3-way Interaction: We are arguing for more than just an amplitude reduction from Recovered to Sedated, which could generate a three-way interaction. One way to see that there is more than just an amplitude effect is to notice in Figure 6 (main body of paper), middle panel, left ERP plot, that the Sedated LDGD (dotted blue line) actually crosses and becomes bigger than Recovered LDGD (filled blue line) exactly when the 3-way interaction is significant at Cz. Such an effect could not arise from a simple “squeezing down” of amplitude towards the GS conditions due to sedation.

S3. Summary of Cluster-level Analysis

Summary Table S.T1 presents the statistics for the main effects and interactions with cluster forming threshold of 0.001, and FWE rate correction of 0.05 (FWE_{corr} = 0.05).

effect	Cluster level analysis		F-value	Peak		
	$P_{FWE-corr}$	K_E		x	y	ms
L	4.77E-15	12216	80.54	9	-20	256
	0.025	780	42.24	68	-43	220
	6.59E-06	3551	35.41	5	26	120
	1.02E-06	4749	33.54	5	20	448
	3.2E-03	1238	30.12	-59	-48	116
G	<E-17	34721	167.74	50	-54	424
	8.88E-16	20081	116.52	5	61	552
S	4.86E-03	900	35.69	-63	-25	108
	1.13E-04	2625	28.68	23	26	116
	1.23E-05	2329	20.87	36	-43	380
LxG	1.75E-11	9372	36.53	-5	9	552
	2.65E-06	3183	26.75	9	-20	256
SxL	2.7E-04	1604	24.41	-5	-14	216
SxG	0.033	604	18.76	-27	-25	340
	0.028	594	17.74	-14	-3	304
SxLxG	0.026	562	19.61	18	-31	420

Table S.T1. The cluster level analysis summary table for all the main effects and interactions, with cluster forming threshold of 0.001 and FWE_{corr} = 0.05. The significant clusters ($P_{FWE-corr}$) and their corresponding cluster size (K_E) are presented for each main effect and interaction. Degrees of freedom = [1, 136]. The peak of each active cluster (F-value) and its location in time and space is also shown. Time is presented relative to the onset of the fifth tone. With regard to spatial position of peaks of clusters, the origin (0,0) is at the electrode Cz and (x,y) shows the coordinates of the peaks in mm; i.e. +x shows a position on right hemisphere and +y a position at the front. Each F-value presents a test at the peak (highest voxel) in a cluster; it reflects the height of this voxel, rather than directly, the significance of the entire cluster. The local effect (L) includes 5 significant clusters, where the peak of the largest cluster appears at 256ms, and has extent (i.e. number of space-time points) KE = 12216. The global effect (G) contains two significant clusters, where the peaks

of the effect appear late, e.g. 424ms. The sedation effect (S) includes 3 clusters. The first two clusters reach their maximum early (i.e. peaks at 108ms and 116ms) and the last cluster reaches its maximum late, at 380ms. The local-global interaction (LxG) has two significant clusters. The larger cluster is maximal in the global effect time region; i.e. 552ms. The smaller cluster is maximal at 256ms during the second phase of the local effect. The sedation-local interaction (SxL) has one active cluster with peaks in an early time region. The sedation-global interaction (SxG) has two significant cluster with peaks at 304 and 340ms. Finally, the sedation-local-global interaction (SxLxG) has a late cluster with a peak at 420ms.

S4. Average Amplitude Variation

We present the statistical results of the analysis of average amplitude variation (tables S.T2 to S.T5) using different window placement strategies:

Source	Df	F	Sig. (i.e. p value)	Partial Eta Squared
Sed	1,17	.551	.468	.031
Lo	1,17	1.214	.286	.067
G	1,17	.058	.812	.003
Sed * Lo	1,17	.336	.570	.019
Sed * G	1,17	.361	.556	.021
Lo * G	1,17	.079	.782	.005
Sed * Lo * G	1,17	2.887	.108	.145

Table S.T2. 2x2x2 Test of within-subjects effects for amplitude variation over [200 end] window at channel Cz. Sed (the variation in P3 amplitude for the main effect of sedation), Lo (the variation in P3 amplitude for the local effect),

G (the variation in P3 amplitude for the global effect), Sed*Lo (the interaction between the sedation and local effects), Sed*G (the interaction between the sedation and global effects), Lo*G (the interaction between the local and global effects), Sed*Lo*G (the interaction between the sedation, local, and global effects).

Source	Df	F	Sig. (i.e. p value)	Partial Eta Squared
Sed	1,17	2.948	.104	.148
Lo	1,17	.952	.343	.053
G	1,17	.142	.711	.008
Sed * Lo	1,17	.961	.341	.054
Sed * G	1,17	.244	.628	.014
Lo * G	1,17	.466	.504	.027
Sed * Lo * G	1,17	3.047	.099	.152

Table S.T3. 2x2x2 Test of within-subjects effects for amplitude variation over [200 end] window at channel Pz. Sed (the variation in P3 amplitude for the main effect of sedation), Lo (the variation in P3 amplitude for the local effect), G (the variation in P3 amplitude for the global effect), Sed*Lo (the interaction between the sedation and local effects), Sed*G (the interaction between the sedation and global effects), Lo*G (the interaction between the local and global effects), Sed*Lo*G (the interaction between the sedation, local, and global effects)

Source	Df	F	Sig. (i.e. p value)	Partial Eta Squared
Sed	1,17	.437	.517	.025
Lo	1,17	5.217	.035	.235
G	1,17	.846	.371	.047
Sed * Lo	1,17	.548	.469	.031
Sed * G	1,17	2.992	.102	.150
Lo * G	1,17	1.294	.271	.071
Sed * Lo * G	1,17	1.204	.288	.066

Table S.T4. 2x2x2 Test of within-subjects effects for amplitude variation at channel Cz. The window for the LD conditions was set over [200 400] and the window for LS conditions was set over [250 end]. Sed (the variation in P3 amplitude for the main effect of sedation), Lo (the variation in P3 amplitude for the local effect), G (the variation in P3 amplitude for the global effect), Sed*Lo (the interaction between the sedation and local effects), Sed*G (the interaction between the sedation and global effects), Lo*G (the interaction between the local and global effects), Sed*Lo*G (the interaction between the sedation, local, and global effects)

Source	Df	F	Sig. (i.e. p value)	Partial Eta Squared
Sed	1,17	2.959	.104	.148
Lo	1,17	4.593	.047	.213
G	1,17	2.335	.145	.121
Sed * Lo	1,17	.072	.791	.004
Sed * G	1,17	.942	.345	.052
Lo * G	1,17	2.832	.111	.143
Sed * Lo * G	1,17	1.402	.253	.076

Table S.T5. 2x2x2 Test of within-subjects effects for amplitude variation at channel Pz. The window for the LD conditions was set over [200 400] and the window for LS conditions was set over [250 end]. Sed (the variation in P3 amplitude for the main effect of sedation), Lo (the variation in P3 amplitude for the local effect), G (the variation in P3 amplitude for the global effect), Sed*Lo (the interaction between the sedation and local effects), Sed*G (the interaction between the sedation and global effects), Lo*G (the interaction between the local and global effects), Sed*Lo*G (the interaction between the sedation, local, and global effects).

This full factorial analysis (tables S.T2 to S.T5) gave no significant effect with the exception of the local effect in tables S.T4 and S.T5 of the Appendix, which was for the varying window approach. This study, however, was focussed on changes in standard deviation of amplitudes associated with the sedation-global interaction. Additionally, one may anticipate a local effect, since the LD and LS conditions employ different windows. The LD conditions use a window (200ms to 400ms) where there is a local effect response, while the LS conditions use a window (250ms to end) where the local response is weakening over time. So, the presence of this local effect is not relevant to our main interest and is also not surprising.

S5. Confidence Intervals, Correction Factor and an Example for Bayesian Analysis

The limits of confidence intervals for all the 8 conditions at channels Cz, and Pz based on the two different window placement strategies are presented in tables S.T6 to S.T9.

	t	Df	Sig. (i.e. p value)	Mean	95% Confidence Interval of the Difference	
					Lower	Upper
RLDGD	14.745	17	.000	5.4148	4.6400	6.1896
RLDGS	10.190	17	.000	5.6029	4.4428	6.7630
RLSGD	10.673	17	.000	5.4983	4.4114	6.5852
RLSGS	14.454	17	.000	5.2258	4.4630	5.9886
SLDGD	11.990	17	.000	5.9079	4.8683	6.9475
SLDGS	12.255	17	.000	5.6565	4.6826	6.6303
SLSGD	14.857	17	.000	5.4901	4.7105	6.2697
SLSGS	10.154	17	.000	5.9609	4.7223	7.1994

Table S.T6. One sample t-test for amplitude variation with 95% confidence interval at channel Cz. Analysis window is from 200ms to the end of the analysed segment. The variation of trial average amplitudes over this window for each condition is used as a participant effect for the one-sample t-test. The mean and confidence interval for each condition are also presented.

	t	df	Sig. (i.e. p value)	Mean	95% Confidence Interval of the Difference	
					Lower	Upper
RLDGD	20.548	17	.000	5.1158	4.5905	5.6411
RLDGS	21.297	17	.000	5.1301	4.6219	5.6383
RLSGD	17.122	17	.000	5.2266	4.5825	5.8706
RLSGS	22.900	17	.000	4.9992	4.5386	5.4598
SLDGD	11.824	17	.000	6.1097	5.0195	7.1999
SLDGS	11.900	17	.000	5.8064	4.7770	6.8359
SLSGD	15.813	17	.000	5.6513	4.8973	6.4053
SLSGS	10.087	17	.000	5.9543	4.7089	7.1998

Table S.T7. One sample t-test for amplitude variation with 95% confidence interval at channel Pz. Analysis window is from 200ms to the end of the analysed segment. The variation of trial average amplitudes over this window for each condition is used as a participant effect for the one-sample t-test. The mean and confidence interval for each condition is also presented.

	t	df	Sig. (i.e. p value)	Mean	95% Confidence Interval of the Difference	
					Lower	Upper
RLDGD	14.578	17	.000	6.1715	5.2783	7.0647
RLDGS	12.364	17	.000	5.7162	4.7408	6.6916
RLSGD	10.438	17	.000	5.6764	4.5291	6.8238
RLSGS	14.263	17	.000	5.3778	4.5823	6.1733
SLDGD	13.013	17	.000	6.3326	5.3059	7.3594
SLDGS	11.739	17	.000	6.0010	4.9225	7.0796
SLSGD	14.877	17	.000	5.6629	4.8598	6.4660
SLSGS	10.206	17	.000	6.1128	4.8492	7.3764

Table S.T8. One sample t-test for amplitude variation with 95% confidence interval at channel Cz. The LD analysis window is from 200ms to 400ms, and the LS window from 250ms to the end of the analysed segments. The variation of trial average amplitudes over the windows for each condition is used as a participant effect for the one-sample t-test. The mean and confidence interval for each condition is also presented.

	t	df	Sig. (i.e. p value)	Mean	95% Confidence Interval of the Difference	
					Lower	Upper
RLDGD	21.651	17	.000	5.7102	5.1538	6.2667
RLDGS	21.292	17	.000	5.2721	4.7496	5.7945
RLSGD	16.672	17	.000	5.3984	4.7152	6.0816
RLSGS	22.680	17	.000	5.1444	4.6658	5.6230
SLDGD	13.180	17	.000	6.5188	5.4753	7.5622
SLDGS	12.496	17	.000	6.0027	4.9892	7.0162
SLSGD	16.040	17	.000	5.8558	5.0856	6.6261
SLSGS	10.112	17	.000	6.1081	4.8337	7.3826

Table S.T9. One sample t-test for amplitude variation with 95% confidence interval at channel Pz. The LD analysis window is from 200ms to 400ms, and the LS window from 250ms to the end of the analysed segment. The variation of trial average amplitudes over the windows for each condition is used as a participant effect for the one-sample t-test. The mean and confidence interval for each condition is also presented.

In order to determine the values of Bayes factors for the sedation-global interaction, we need to determine the range that the priors used in this analysis should take. As explained in the “Single Trial Analysis” subsection of Materials and Methods section for prior estimation, we first calculate the limits of the 95% confidence interval for all the conditions (i.e. RLDGD, RLDGS, RLSGD, RLSGS, SLDGD, SLDGS, SLSGD, SLSGS), to find the limits of the uniform prior. Using the t-distribution, tables S.T6 to S.T9 provide details for confidence intervals and average effects for each condition at channels Cz, and Pz, based on the two different window placement strategies.

To give an example of how this calculation is made, we take the confidence interval information from S.T6 for channel Cz and a window from 800ms to the end of the analysed segment. To find the maximal interaction in equation (1) in the main body of this paper, components RLDGD, RLSGD, SLDGS, and SLSGS of the interaction take the upper limits and RLDGS, RLSGS, SLDGD, and SLSGD take the lower limits of the respective confidence intervals. Then, the maximal interaction is (6.1896 + 6.5852 + 6.6303 + 7.1994 - 4.4428 - 4.4630 - 4.8683 - 4.7105)/2 = 4.0599. By the same token, the interaction is minimal when RLDGD, RLSGD, SLDGS, and SLSGS take the lower limits and RLDGS, RLSGS, SLDGD, and SLSGD take the upper limits of corresponding confidence intervals, i.e. the minimal interaction is (4.6400 + 4.4114 + 4.6826 + 4.7223 - 6.7630 - 5.9886 - 6.9475 - 6.2697)/2 = -3.7562. The maximum of the absolute values of -3.7562 and 4.0599 is 4.0599, so prior limits are then set at [-4.0599 4.0599], or approximated by [-4 4]. The next step is to find the mean and standard deviation of the likelihood distribution. Using equation (1) and table S.T6 for the mean of each condition, the mean difference of the SxG interaction is (5.4148 + 5.4983 - 5.6029 - 5.2258 - 5.9079 - 5.4901 + 5.6565 + 5.9609)/2 = 0.1519. Furthermore, looking at summary Table 1 for channel Cz, and strategy 1, the F-value

is 0.361 and $t = \sqrt{F} = 0.6$. Thus the standard deviation of the distribution is mean difference/t = 0.1519/0.6 = 0.2530. This should be corrected for degrees of freedom smaller than 30 (df = 17) with the correction factor of 1+20/(17*17) = 1.0692; i.e. SD = 1.0692x0.2530 = 0.2705.

S6 Composition of Conditions

In our analysis, we have kept all trials in all conditions that passed artefact rejection. Keeping trial counts high in this way, enabled us to obtain the most accurate estimates we can for ERPs in each condition. A consequence of this is though that trial counts are not equalised across conditions. Unbalanced conditions of this kind are common in ERP research, and not a problem in and of themselves, as long as trial counts in all condition-participant bins are kept high, which they are here (in the smaller GD conditions, all participants had more than 19 trials in each condition, with an average of 31 and standard deviation of 8.28). However, there has been evidence reported of strategic effects unfolding through blocks in the local-global task (Chennu et al 2013). Specifically, a Contingent Negative Variation (CNV) was identified (Chennu et al 2013), which may be expected to manifest differently at different points during a block. If this were the case, we may see a frontally negative trend that has a different steepness for trials of different conditions. This is because there are temporal regularities in blocks, such as, global deviant trials typically occurring after long sequences of global standard trials. We considered whether such a trend could confound our data and concluded that it could not for the following reasons.

1) Figure 5 top panel shows ERPs of the four constituent conditions of the Local x Global interaction. We see a trend in the data that is apparent in the LSGS and LDGS conditions. However, importantly, the trend is the same in the two

conditions. Accordingly, it seems very unlikely that the two different block types (LS as GS, i.e. XXXXX, and LD as GS, i.e. XXXXY) exhibit different trend patterns. This is important, since it also suggests that, if there is a trend present, it would be the same in LDGD and LSGD trials, since such trials appear with the same presentation criteria (e.g. frequency) in LS as GS and LD as GS blocks, respectively.

This would then mean that any trend that might be present would be subtracted out in the interaction contrast, which can be written as (LDGD – LSGD) – (LDGS – LSGS), where both conditions of the first difference would have the same trend, as would both conditions of the second.

2) The reasoning of the previous point carries over to the following conditions: Local, Sedation and Local x Sedation, since the number of GS and GD trials is equalised in all constituent conditions of each of these contrasts.

3) The Global effect is where a problem would be most likely to arise, since it directly contrasts GDs (infrequent) with GSs (frequent). If there was an issue though, it should be observable in our time series before the global effect starts. However, considering Figure 4, ERPs in G panels, one can see that the GD and GS time series are well aligned from 200ms before the 5th tone to 200ms after it. Certainly, there does not seem to be any evidence of a differential trend in the two conditions in this 400ms time period.

4) The absence of a differential trend at relevant electrodes for the Global effect suggests that our findings in the Sedation x Global contrast should also not be confounded by differential trends. This is because the Sedation x Global

contrast is a difference of two GD minus GS terms, neither of which would be expected to be confounded.

5) Finally, confound-freeness of the Local x Global interaction (as argued for under point 1 above), would carry over to the L x G x S interaction, which is a difference of two such Local x Global terms.

S7 Standard error of mean plots for recovered and sedation ERPs

The shaded areas in figures S3 to S6 present error bar for each corresponding ERP, where the standard error of the mean is depicted relative to the mean (i.e. ERP) at each time point. The standard error of mean is $\frac{\sigma}{\sqrt{N}}$, where *N=18* is the sample size (i.e. 18 participants) and σ is the standard deviation of the mean distribution (or ERP distribution) at each time point. Therefore, the error bar at each time point sits in between $\mu - \frac{\sigma}{2*\sqrt{18}}$ and $\mu + \frac{\sigma}{2*\sqrt{18}}$, where μ is the grand mean ERP value at a time point.

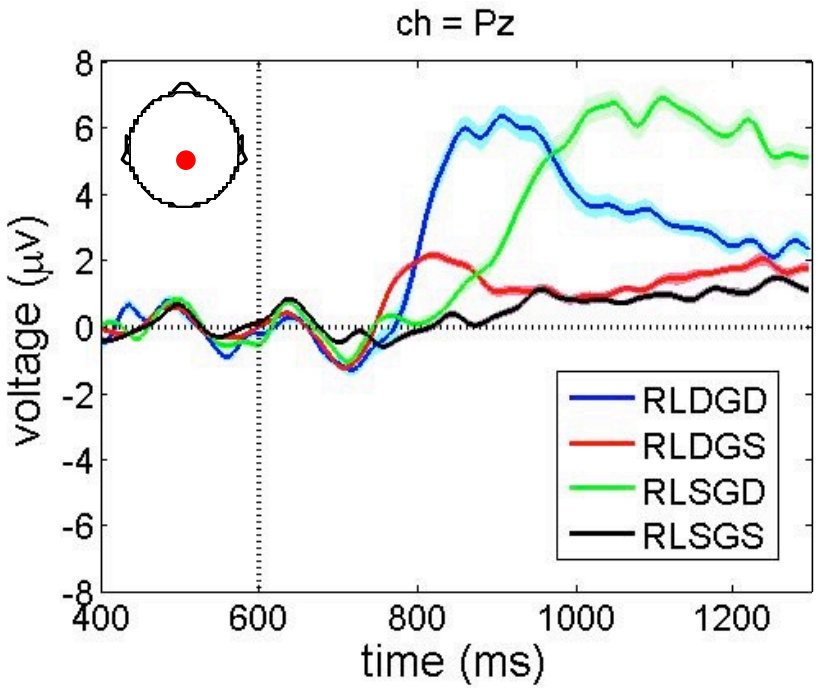


Figure S3. Representation of the standard error of mean alongside each recovered ERP condition at channel Pz.

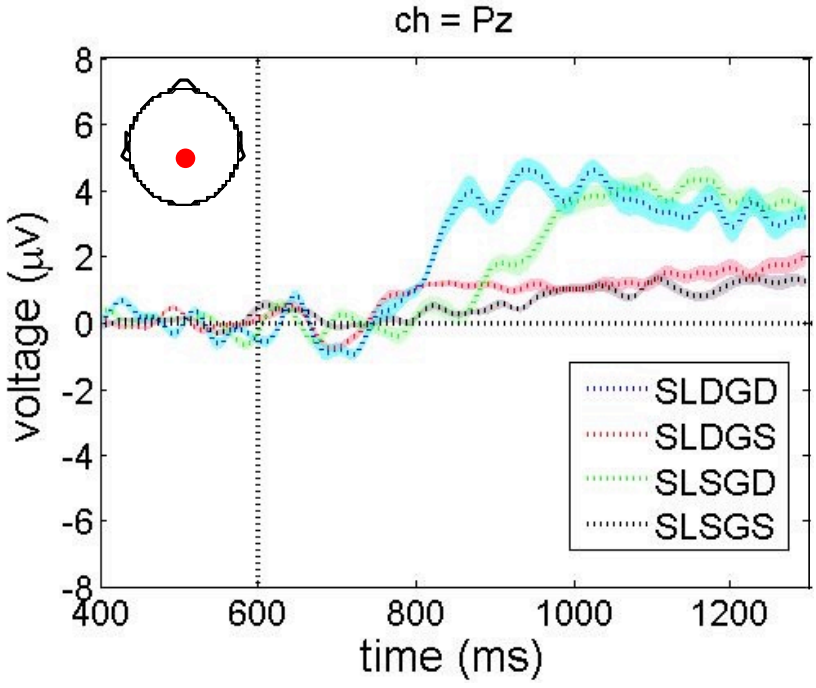


Figure S4. Representation of the standard error of mean alongside each sedated ERP condition at channel Pz.

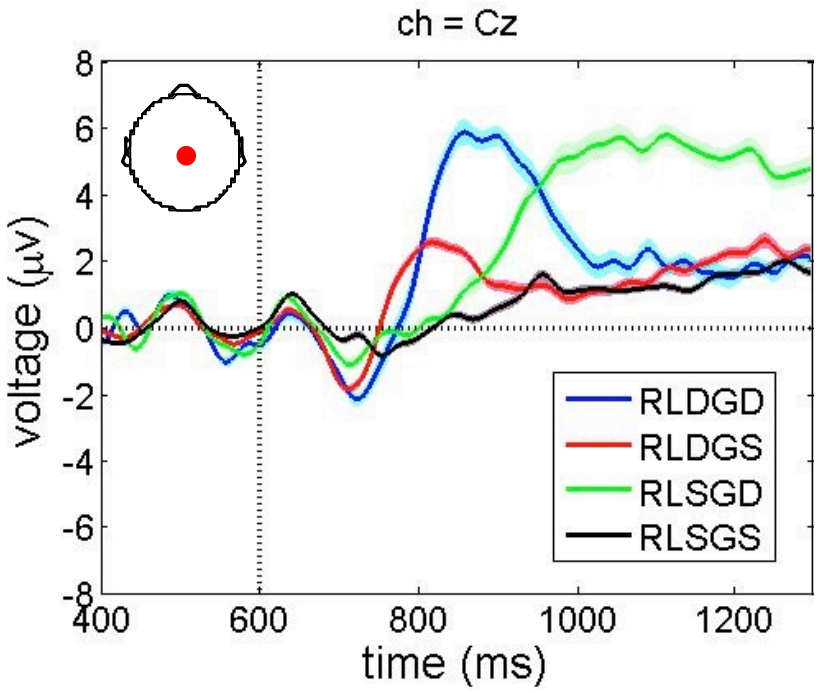


Figure S5. Representation of the standard error of mean alongside each recovered ERP condition at channel Cz.

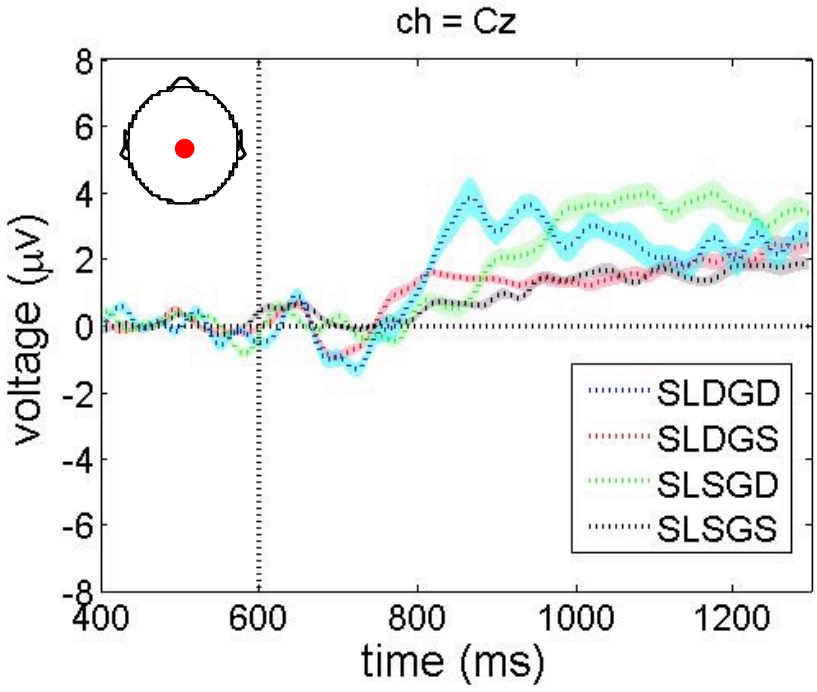


Figure S6. Representation of the standard error of mean alongside each sedated ERP condition at channel Cz

S8. A single-Subject ERP plots

Here, we show an example of ERP plots for a single subject (c.f. figures S7 to S10). As observed, the ERP responses are sharper compared to the average responses in figure 6, across all the participants.

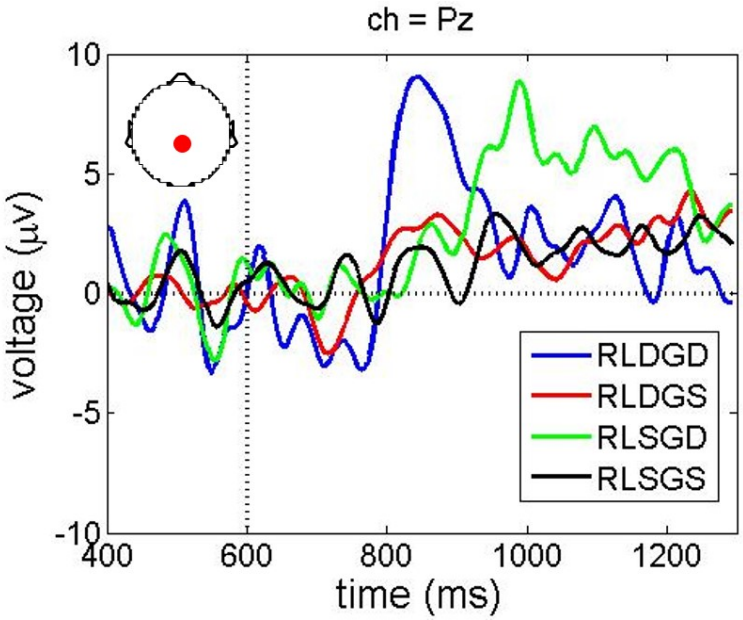


Figure S7. ERP plots for a single subject and recovered conditions at channel Pz.

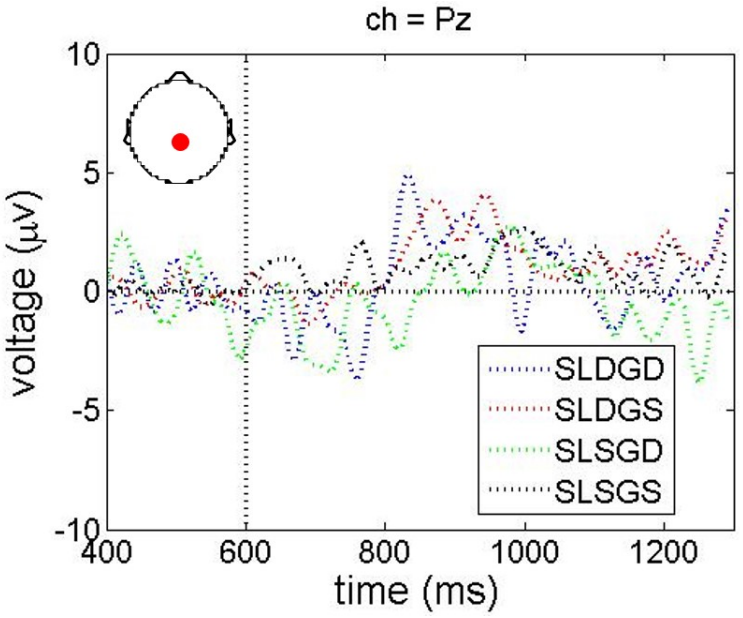


Figure S8. ERP plots for a single subject and sedated conditions at channel Pz.

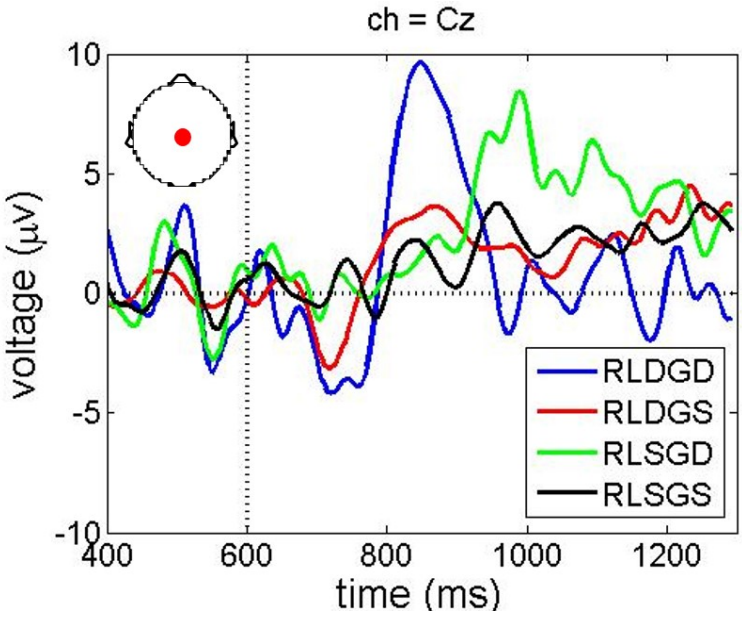


Figure S9. ERP plots for a single subject and recovered conditions at channel Cz.

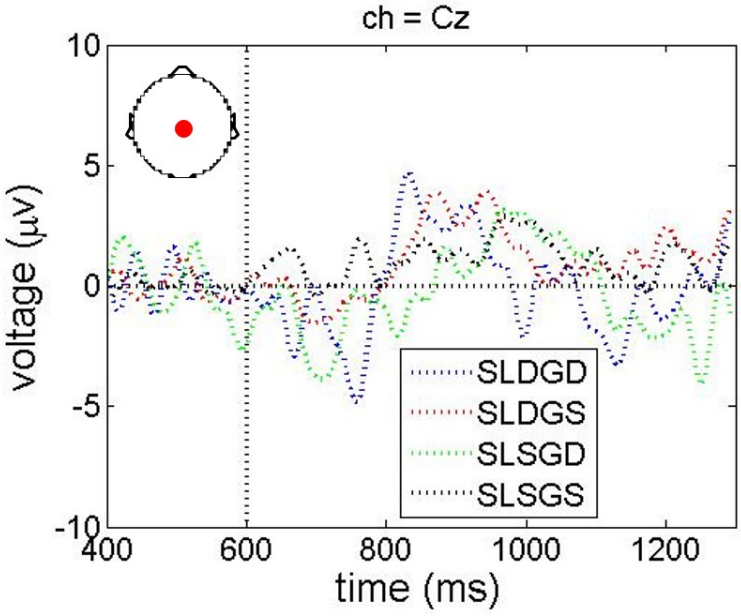


Figure S10. ERP plots for a single subject and sedated conditions at channel Cz.

S9. The base line period scalp maps:

Figure S11 represents the scalp maps for the baseline period. The baseline period starts 200ms before the onset of the fifth tone. As expected, the average of scalp maps over time is almost zero over the base line period. Overall, no active cluster extended over time-space is observed.

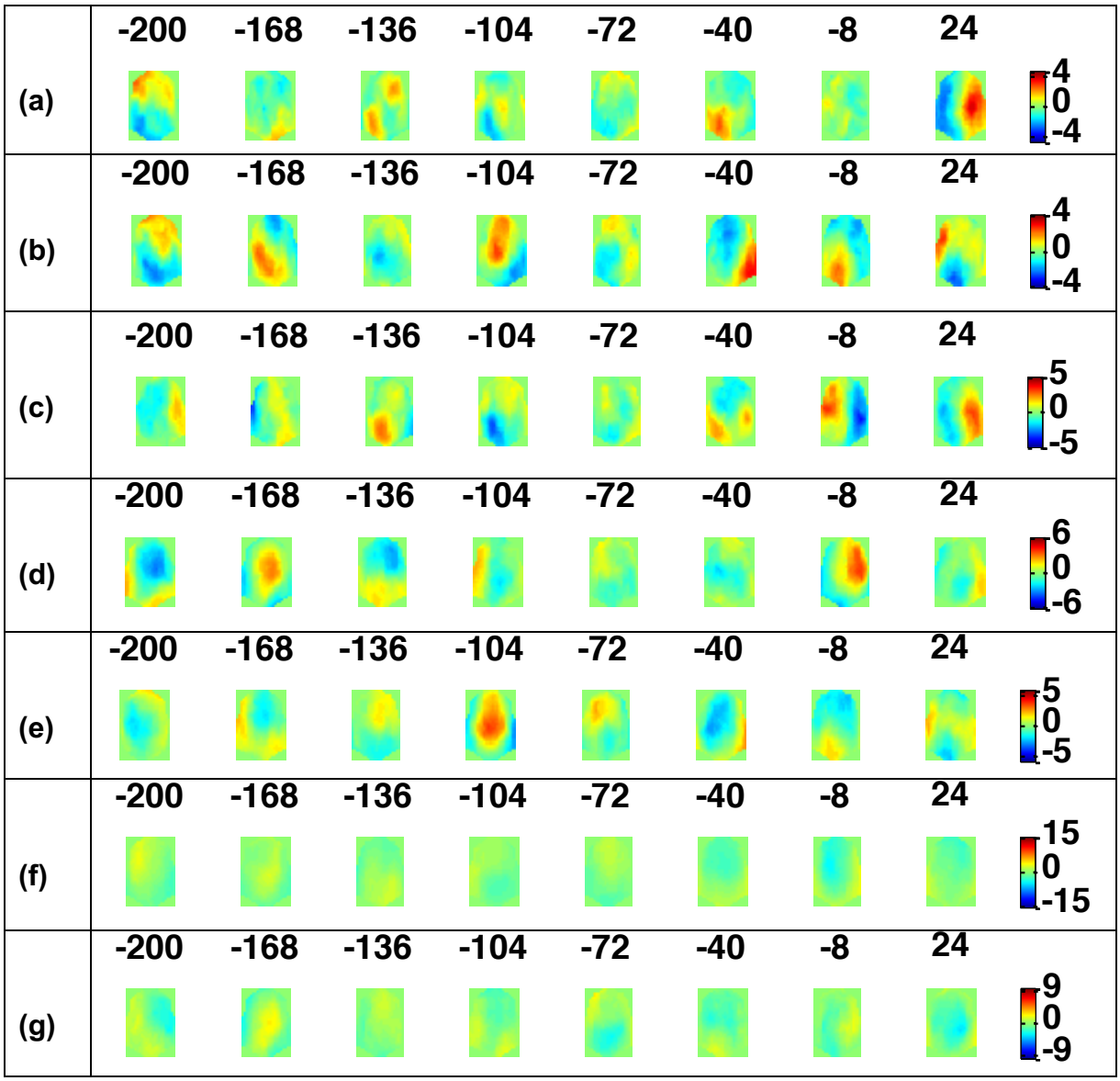


Figure S11. The baseline period for the three-way interaction (a), the sedation-global interaction (b), the sedation-local interaction (c), the local-global interaction (d), the sedation effect (e), the global effect (f), and the local effect (g).