



## Investigating Established EEG Parameter During Real-World Driving

Janna Protzak<sup>1\*</sup> and Klaus Gramann<sup>2,3,4</sup>

<sup>1</sup> Junior Research Group FANS (Pedestrian Assistance System for Older Road User), Institute of Psychology and Ergonomics, Technische Universität Berlin, Berlin, Germany, <sup>2</sup> Biological Psychology and Neuroergonomics, Technische Universität Berlin, Berlin, Germany, 3 Center for Advanced Neurological Engineering, University of California, San Diego, San Diego, CA, United States, <sup>4</sup> School of Software, University of Technology, Sydney, NSW, Australia

In real life, behavior is influenced by dynamically changing contextual factors and is rarely limited to simple tasks and binary choices. For a meaningful interpretation of brain dynamics underlying more natural cognitive processing in active humans, ecologically valid test scenarios are essential. To understand whether brain dynamics in restricted artificial lab settings reflect the neural activity in complex natural environments, we systematically tested the auditory event-related P300 in both settings. We developed an integrative approach comprising an initial P300-study in a highly controlled laboratory set-up and a subsequent validation within a realistic driving scenario. Using a simulated dialog with a speech-based input system, increased P300 amplitudes reflected processing of infrequent and incorrect auditory feedback events in both the laboratory setting and the real world setup. Environmental noise and movement-related activity in the car driving scenario led to higher data rejection rates but revealed comparable theta and alpha frequency band pattern. Our results demonstrate the possibility to investigate cognitive functions like context updating in highly artifact prone driving scenarios and encourage the consideration of more realistic task settings in prospective brain imaging approaches.

Keywords: MoBI, real driving, P300, electroencephalography (EEG), auditory feedback

## **1. INTRODUCTION**

To improve our understanding of human cognition and the underlying brain dynamic processes in real life situations, ecological task settings are needed that allow complex and realistic behaviors (Engel et al., 2013). On-road driving scenarios are such an example of an ecological task setting in which, in contrast to simulated driving, incorrect behavior can have drastic consequences. While laboratory studies allow controlled investigations of specific cognitive and behavioral processes, it is not clear whether these phenomena can be observed in real life conditions. This is especially the case for behaviors that involve active movements of participants which provide sensory feedback that itself influences brain dynamics and cognition (e.g., Gramann, 2013). However, the application of established brain imaging methods like electroencephalography (EEG) in more natural task settings are hindered by artifacts induced by active behavior. Non-brain activity like muscle and eye movements, or electric and mechanical artifacts can severely impact the signal quality on the sensor level. However, advances in mobile amplifier systems and developments in data analyses approaches can overcome these problems. The recently developed Mobile Brain-Body Imaging (MoBI) approach (e.g., Makeig et al., 2009; Gramann et al., 2011, 2014) overcomes the restrictions of traditional imaging modalities by using ambulatory EEG or NIRS devices combined

**OPEN ACCESS** 

#### Edited by:

Claudia Gianelli, Universität Potsdam, Germany

#### Reviewed by:

Edmund Wascher, Leibniz-Institut für Arbeitsforschung an der TU Dortmund (IfADo), Germany Sven Hoffmann. German Sport University Cologne, Germanv

> \*Correspondence: Janna Protzak

janna.protzak@tu-berlin.de

#### Specialty section:

This article was submitted to Cognition, a section of the iournal Frontiers in Psychology

Received: 07 July 2018 Accepted: 02 November 2018 Published: 23 November 2018

#### Citation:

Protzak J and Gramann K (2018) Investigating Established EEG Parameter During Real-World Driving. Front. Psychol. 9:2289. doi: 10.3389/fpsyg.2018.02289

1

with motion capture and other data streams that allow active behavior (e.g., Gwin et al., 2010; Jungnickel and Gramann, 2016; Banaei et al., 2017). MoBI studies demonstrate that brain activity can be distinguished from environmental and behavioral artifacts, opening up new possibilities for more realistic test and acquisition scenarios outside restricted laboratory set-ups. Driving a car is one such realistic scenario that is highly relevant for a large part of the population but represents a hostile recording environment for EEG recordings. Driving takes place in non-shielded environments with electronic equipment surrounding the driver and the task requires complex behaviors, including movement of the eyes, the head, as well as the arms and shoulders, that are typically restricted in standard laboratory settings to avoid movement-related artifacts from distorting the signal of interest. Analyzing human brain dynamics in a real driving scenario can thus be considered a stress test for comparison of EEG parameters, e.g., event-related potentials (ERP), obtained during real-world driving with parameters established in traditional laboratory settings including car simulators. If established parameters like the event-related P300 component can be replicated in real driving scenarios, EEGdata can be used to improve our understanding of how drivers process information while controlling a vehicle in a realistic environment. Providing direct access to the driver's neuronal responses during different driving process phases, EEG might serve the development and evaluation of user centered designs for technical assistance systems in the safety-critical driving environment (e.g., Brouwer et al., 2017).

So far, only a few studies have recorded and analyzed brain activity in real-life driving tasks and the majority of these studies focus on workload measures (Kohlmorgen et al., 2007) or vigilance (e.g., Kecklund and Akerstedt, 1993; Papadelis et al., 2007; Schmidt et al., 2009; Simon et al., 2011; Sonnleitner et al., 2014). Haufe et al. (2014) present results from a driving study for an automated braking assistance system using EEG and EMG data demonstrating the potential use of event-related potentials (ERP) to enhance automated driving technology. Because the focus of the study by Haufe and colleagues was on the replication of classification results from an earlier driving simulator study (Haufe et al., 2011), no quantitative analyses of ERP components were provided. Zhang et al. (2015) executed a combined simulator and real car study to develop a brain-computer interface (BCI) for detecting error-related EEGactivity. Despite a clear focus on classification accuracies and a small sample size for the real car experiment, the ERP results revealed comparable patterns for both acquisition scenarios, even though these were not specifically addressed in the discussion. Krol et al. (2017) investigated a BCI approach during interaction with an automated cruise control system in a real driving scenario. The authors demonstrate high classification accuracies for unexpected events during cruise control. However, as the focus was on classification and not replication of specific EEG features, no general conclusion can be drawn from this study about the replicability of established EEG parameters.

As no previous study has provided a detailed analysis of event related potentials during real life driving, it is still an open question whether systematic ERP-analysis is possible with

data recorded in real driving scenarios and whether the results can be compared with those from traditional laboratory EEG recordings. We addressed this question by comparing the eventrelated P300 recorded during a dual-task driving scenario and within a highly controlled single task laboratory setup. The auditory secondary task consisted of an interaction of the participant with a speech input device, resembling a common on-road secondary task. ERPs with onset of incorrect feedback from the speech input device were analyzed with a focus on the event-related P300 component, a positive deflection in the ERP that represents a well-established parameter for analyzing cognitive functions like attention and memory, substantiated by results from extensive laboratory assessments with numerous and heterogeneous groups of persons (Sutton et al., 1965, for reviews see Fabiani et al., 1987; Picton, 1992). Increased P300 amplitudes can be observed for infrequent targets in a stream of frequent stimuli (for a review see Polich, 2007), for random and time varying targets in single-stimulus paradigms (Polich et al., 1994), for unexpected feedback events (Horst et al., 1980) or when erroneous actions are observed (de Bruijn et al., 2007). It has been argued that the reversed relationship of stimulus probability and P300 amplitudes indexes the amount of working memory updating after deviant events that is necessary for the processing of the preceding stimulus (Donchin et al., 1978; Donchin and Coles, 1988) and that the P300 mediates between stimulus and response processes (Verleger et al., 2005). Furthermore, dualtask studies have shown that demanding primary tasks can result in reduced P300 amplitudes evoked by secondary task stimuli. This amplitude reduction was interpreted as reflecting resource reallocation processes between parallel executed tasks (Isreal et al., 1980a,b; Sirevaag et al., 1989).

The P300 was expected to reflect processing of infrequent erroneous auditory feedback events in both recording environments with adequate data preprocessing in the real driving setup. Specifically, higher P300 amplitudes were expected for rare incorrect feedback events compared to correct feedback trials. Furthermore, modulations in P300 amplitudes in the driving task might index the amount of processing resources that are needed to perform the driving task. In addition, the baseline EEG power spectra from both recordings were analyzed to examine possible tonic differences and to distinguish them from phasic event-related effects.

## 2. STUDY 1: LABORATORY SETUP

## 2.1. Method

#### 2.1.1. Participants

Eighteen participants volunteered for the first study. Three data sets had to be discarded due to extensive artifacts in the EEG data. The analyzed sample included 15 healthy adults (10 female, 20–35 years of age, mean 28 years). All volunteers were right handed as assessed by a German adaptation of Edinburgh handedness inventory (Oldfield, 1971) and none reported a history of neurological problems. The study was carried out in accordance with recommendations of the German Psychological Society and all participants gave written informed consent in accordance with the Declaration of Helsinki. At the time of the

data recordings, ethics approvals were not required by guidelines of the Technische Univeristät Berlin.

#### 2.1.2. Experimental Design and Procedure

Participants were seated in front of a 19" screen for visual stimulus presentation with their index fingers positioned on the marked ctrl-buttons on a standard keyboard on a table in front of them. Auditory feedback was presented through speakers placed at either side of the screen. A pool of common German first names with at least two syllables served as the stimulus material. All names were digitized as auditory feedback cues with Natural Reading Software (Natural Reading Software, Vancouver, BC Canada) and used for a simulated dialog between the driver and a technical speech based input system.

Each trial started with a black and gray flashing display for 800 ms, followed by a grayscreen for 200 ms (Figure 1). Three randomly chosen names from a pool of 145 forenames were presented consecutively in black letters on a gray background for 2,000 ms each. In parallel, the same names were read aloud in their digitized version by a synthesized female voice. Participants were asked to remember all three names and then speak out loud the name of the sequence position that was randomly displayed at the end of the trial (e.g., "two back" indicating to repeat the second name). A subsequent response interval lasted for 5,000 ms followed by an auditory repetition of the participant's response. In 80% of all cases the auditory feedback matched the stated name (eg. "Ella"), while in 20 % of all cases, only the last syllable (eg. "la") was replayed. Correct and incorrect feedback trials were randomly presented in each trial sequence. Participants were required to wait for a tone after another 1,000 ms to categorize the feedback. Correct repetitions had to be confirmed by a button press with the right index finger on the right ctrl-key and incorrect repetitions had to be indicated by pressing the left crtl-key using the left index finger.

The task protocol followed a Wizard of Oz procedure where people believed to interact with a technical system even though operations were at least partially controlled by a human operator (cf. Dahlbäck et al., 1993). In the present case, the participants' spoken responses were not analyzed by an automated speech recognition system but by the experimenter. The manually registered response was transferred to the experimental program to implement a fixed and random error rate in the auditory feedback. Subsequent interviews revealed that none of the participants recognized the manipulation. The study consisted of six blocks of 50 trials each. The entire procedure took 2.5 h on average.

#### 2.1.3. EEG-Recording and Pre-processing

EEG-data were recorded continuously from 64 active electrodes (Brainproducts GmbH, Gilching, Germany), mounted in an elastic cap according to the extended international 10–20 system (Chatrian et al., 1985), with the exception of positions PO7 and PO8, which were placed below the left and right eye, respectively, to measure electroocular activity. The data were digitized with a sampling rate of 1,000 Hz. Prior to data recordings, impedances were brought below 5 k $\Omega$ . Off-line preprocessing and data analysis were performed in Matlab 2015

(MATLAB, The MathWorks Inc., Natick, MA, USA), using Eeglab-based routines (Delorme and Makeig, 2004). The data were filtered with a 0.1 to 100 Hz band pass filter and the sampling rate was subsequently reduced to 500 Hz. Artifact contaminated channels (M = 11, SD = 3.5) were removed using automatic rejection (5 standard deviations of the mean kurtosis value or 3 standard deviations from mean probability distribution of each single channel) and subsequent manual visual inspection. Afterwards, all channels were re-referenced to an average reference calculated by the remaining channels. At this point, two copies were made of each data set. The first set was filtered with a 1 Hz high pass filter and only used for independent component analysis (ICA). The second set was filtered with a 40 Hz low pass filter and used for any further analysis. Spatially static and maximally temporally independent components (ICs) were calculated for each participant on the first set using adaptive mixture independent component analysis algorithm (AMICA, Palmer et al., 2008) which allows flexible source modeling for each component by using Generalized Gaussian density models. The applied AMICA settings included that one model was trained, three base components densities were assumed for the mixture models and the number of rejections of unlikely data samples was set to three. The resulting ICs weighs were mapped on the 40 Hz low pass filtered sets for the ERP analysis. ICs representing eye movements were categorized for each participant (M = 3, SD = 0.6) by means of scalp maps and activation time courses. Eye movement activity was removed from the recordings by removing ocular ICs and subsequent back-projection to the sensor level.

All resulting data sets were segmented to 1,800 ms epochs, starting 300 ms before the onset of the auditory feedback. For each participant, epochs were automatically discarded if amplitudes exceeded  $+/-80 \ \mu$ V or if the measured probability of a trial exceeded a criterion of 6 standard deviations of the mean calculated probability distribution on a single channel level or 3 standard deviations for all channels. In total, 2,604 correct feedback trials (M = 174, SD = 29.9) and 656 incorrect feedback trials (M = 44, SD = 6.1) were considered for the analysis.

#### 2.1.4. Data Analysis

Averaged correct and incorrect feedback amplitudes were analyzed relative to a 300 ms pre-stimulus baseline (300 -0 ms before feedback onset). The P300 time windows and electrode sites (Fz, Cz, Pz) for analysis were selected based on the literature (e.g., Johnson, 1993) and visual inspection of the grand averages. A 100 ms -time window around the most positive peak at parietal electrode site Pz (738 - 838 ms after stimulus onset) was chosen for P300-analysis. Mean P300 amplitudes were assessed by  $2 \times 3$  repeated measures of variance (ANOVA) with the factors feedback type (correct vs. incorrect) and electrode site (Fz, Cz, Pz). Degrees of freedom were adjusted by means of the Greenhouse-Geisser method in case of deviations from sphericicity. Post-hoc t-tests were calculated for each condition at each electrode to evaluate differences in the topographical distribution of the measured activations and tested against correspondent Bonferroni-corrected alpha levels.



## 2.2. Results

Stimulus-locked ERP-waveforms for incorrect and correct feedback are shown in **Figure 2**. ANOVA results for the main P300 peak time window revealed significant main effects for feedback type,  $F_{(1,14)} = 59.93$ , p < 0.001,  $\eta_p^2 = 0.81$  and electrode site,  $F_{(1.24,17.29)} = 22.11$ , p < 0.001,  $\eta_p^2 = 0.61$ . Mean P300 amplitudes were significantly higher for incorrect ( $M = 2.21\mu$ V,  $SD = 1.05\mu$ V) as compared to correct feedback ( $M = 0.07\mu$ V,  $SD = 0.96\mu$ V). Activity for both feedback conditions increased from frontal electrode site Fz ( $M = -0.99\mu$ V,  $SD = 2.23\mu$ V) toward more posterior sites Cz ( $M = 1.45\mu$ V,  $SD = 1.23\mu$ V),  $t_{(14)} = -5.88$ , p < 0.001, and Pz ( $M = 2.97\mu$ V,  $SD = 0.91\mu$ V),  $t_{(14)} = -5.23$ , p < 0.001. A further increase was measured from electrode site Cz to Pz,  $t_{(14)} = -3.02$ , p = 0.009.

A significant electrode × feedback type interaction,  $F_{(2,28)} = 14.65$ , p < 0.001,  $\eta_p^2 = 0.51$ , reflected amplitude differences between correct feedback trials with lower values at Fz ( $M = -0.96\mu$ V,  $SD = 2.06\mu$ V) compared to Cz ( $M = 0.53\mu$ V,  $SD = 1.44\mu$ V), t(14) = -4.41, p = 0.001, and compared to Pz ( $M = 1.29\mu$ V,  $SD = 0.62\mu$ V),  $t_{(14)} = -3.63$ , p < 0.003. Incorrect feedback trials resulted in reduced activity at Fz ( $M = -0.71\mu$ V,  $SD = 3.01\mu$ V) compared to Cz ( $M = 3.17\mu$ V,  $SD = 1.70\mu$ V),  $t_{(14)} = -6.08$ , p < 0.001, and Pz ( $M = 4.85\mu$ V,  $SD = 1.59\mu$ V),  $t_{(14)} = -5.08$ , p < 0.001. Furthermore, incorrect feedback elicited larger amplitudes than correct feedback at Cz,  $t_{(14)} = -6.49$ , p < 0.001, and Pz,  $t_{(14)} = -9.81$ , p < 0.001, but not a frontal site Fz,  $t_{(14)} = -0.44$ , p = 0.665.

## 2.3. Discussion Study 1

For the laboratory study we used a well-controlled experimental setup to establish a baseline for the experimental manipulation in the subsequent driving task. The analysis focused on the sensitivity of the P300 as an index for the processing of improbable and erroneous events. As expected, the task manipulation elicited differences in event-related brain activity with increased P300 amplitudes for the infrequent incorrect feedback trials. The analysis revealed a posterior distribution with most pronounced differences between correct and incorrect feedback trials over parietal sites. This activation was absent in trials containing correct feedback information. Similarly, studies have shown the P300 amplitude to be sensitive to the subjective probability of an event (Horst et al., 1980) and to errors in picture sequences (de Bruijn et al., 2007). In our case, the P300 appears

to reflect enhanced processing costs for the categorization of the less frequent and unintended erroneous feedback events. This is in line with interpretations of the functionality of the P300 that claim that the P300 reflects the context updating within the evaluation process of new events (Donchin and Coles, 1988). In the present tasks, participants expected to hear a repeat of their own speech input. Consequently, the large P300 for incorrect fragmented feedback most likely displayed the memory update after the mismatch between the anticipated and received feedback. Moreover, as the less often incorrect feedback required a different manual button press, deviating response requirements might also be depicted by these changes (Verleger et al., 2005). The results from Study 1, confirmed our approach for investigating P300 activity for rare and deviant auditory feedback. Consequently, the procedure was applied in the following in-car recordings.

## 3. STUDY 2: DRIVING SETUP

Study 2 was conducted in a real driving setting to test whether human brain dynamics reflective of deviance detection can be recorded while participants actively drive a car. The same task as in the laboratory recordings was used to allow a direct comparison. Data processing procedures were guided by laboratory study routines reported in section 2.1.

## 3.1. Method

#### 3.1.1. Participants

Seventeen participants volunteered in the second study. Data from one participant had to be excluded from analysis due to technical problems during data recording, and data from a second participant had to be removed due to insufficient data quality. The analyzed sample included 15 adults (10 female, 22– 36 years of age, mean 28 years,). All participants held a valid driver license for at least 2 years. As in Study 1, all volunteers complied with the requirements and were tested under the same conditions. None had participated in Study 1.

#### 3.1.2. Experimental Design and Procedure

Participants performed the same audio feedback task with identical stimulus material and time course as described in Study 1 (see section 2.1.2). Only set-up modifications for the in-car realization are described here. The driving tests took place on a part of a restricted runway (length: approx. 1.5 km) of a



former military airfield in Brandenburg, Germany (**Figure 3**). A gear shift *Volkswagen Touran* was provided as test vehicle by the Department of Human-Machine Systems, TU Berlin. Audio feedback was transmitted through portable speakers located in the front interior. Names were presented on a 7.6" TFT-display, mounted on the central console. Two buttons were added to the steering wheel, in a convenient position that allowed for safe steering and button presses with the left and right thumb.

A blocked 1.2 km section of the runway served as the test track (Figure 3). All participants had time to familiarize with the car before maneuvering the car to the starting position at the head of the test track. Test blocks were defined by driving the test track twice back and forth (= 4.8 km). Participants were instructed to accelerate the car to  $40 \ km/h$  and to shift into the fourth gear at the beginning of each run. Speed and gear had to be maintained until the end of the test track was reached (indicated by a pylon). Behind that point, the car had to be turned around and accelerate again for the next test run in the reversed direction. Throughout each block, the experimental task was only started when speed was within a range of 40 km/h + / - 3 km/h (monitored via Control Area Network Data). For economic reasons and to keep up alertness, task blocks alternated with blocks in which participants worked on an acceleration and braking task, not reported here. The total number of completed test blocks differed individually (range 12-14 blocks and 80-120 trials) dependent on weather and the participant's individual condition.



#### 3.1.3. EEG Recording and Preprocessing

The EEG recording setup and preprocessing steps followed the protocol for the laboratory recordings. Data were recorded with 64 active electrodes digitized with a sampling rate of 500 Hz. Impedances were kept below 5 k $\Omega$ . All data sets were offline filtered with a high-pass filter of 0.1 Hz and a low-pass filter of 40 Hz. Again, after automatic and visual inspection artifact contaminated channels were discarded (M = 9, SD = 3.1) and the remaining channels were re-referenced to an average reference. As in Study 1, two copies were made of each

accordingly preprocessed data set. The first set was filtered with a 1 Hz high pass filter and only used for independent component analysis. The second data set was filtered with a 40 Hz low pass filter and used for any further reported analysis. The calculated IC weigths were map on the 40 Hz low pass filtered sets and ICs representing eye movements (M = 4, SD = 1.1) were removed. The resulting data were back projected to the channel level. Trials from the epoched data sets were automatically rejected if any channel contained amplitudes that exceeded  $+/-80 \,\mu$ V. Slightly broader probability criterions (6 SD on single channel level and 3 SD for all channels) were applied for the automated rejection based on deviation from the mean probability distribution to adapt to the generally more fluctuating data quality of the in-car recordings. In sum, 1,222 correct trials (M = 81, SD = 20.0) and 296 incorrect trials (M = 20, SD = 5.5) were considered for analysis.

#### 3.1.4. Data Analysis

Activity at midline electrodes Fz, Cz, and Pz were averaged for correct and incorrect feedback trials and, respectively, calculated in relation to a 300 ms baseline time window preceding the auditory feedback onset. For the analysis of amplitude differences the 100 ms time widow (852 – 952 ms) around parietal peak activity was specified. Mean amplitudes in the P300 time-window were subjected to a  $2 \times 3$  ANOVA with the factors feedback type (correct vs. incorrect) and electrode site. Greenhouse-Geisser corrections were applied and Bonferroni-corrected *t*-tests were calculated for *post-hoc* comparisons of factor levels.

#### 3.2. Results

The analyses of mean amplitude values revealed a main effect for the factor feedback type,  $F_{(1,14)} = 31.67$ , p < 0.001,  $\eta_p^2 =$ 0.69, with higher amplitudes for incorrect feedback (M = $2.04\mu$ V,  $SD = 1.81\mu$ V) compared to correct feedback (M = $-0.75\mu$ V,  $SD = 1.08\mu$ V). The main effect for electrode site,  $F_{(1.40,19.61)} = 2.55$ , p = 0.117,  $\eta_p^2 = 0.15$ , and the interaction effect of feedback type × electrode site,  $F_{(2,28)} = 0.84$ , p =0.444,  $\eta_p^2 = 0.06$ , were not significant (**Figure 4**).

## 3.3. Discussion Study 2

In Study 2, we tested whether the results from Study 1 could be replicated when the identical task had to be accomplished during real driving. As in the laboratory assessment, incorrect feedback elicited larger amplitudes in the P300 time window compared to correct feedback. Although ongoing parallel cognitive and motor processes are needed to solve the driving task, differences in neural response patterns for regularities and discrepancies in auditory feedback could be replicated.

In contrast to Study 1, no significant topographic variations in P300 amplitudes over midline electrodes were found. This activity pattern might be explained by the enhanced complexity of the driving task. Frontal P300 activity as an index of an orienting response has been reported to be dependent on time on task and to diminish with habituation (e.g., Courchesne, 1978, for a review see Friedman et al., 2001). However, less pronounced reductions in frontal activity were found for more complex tasks (Segalowitz et al., 2001). The broad grand average waveform activation pattern including frontal activity in the driving scenario might be due to the fact that the driving task counteracted habituation effects in the secondary task. As the driving task required constant attention, fewer resources might have been available for obtaining automated processes in the auditory secondary task.

# 4. COMPARISON OF RECORDING ENVIRONMENTS

For comparison with the data recorded in Study 1, additional data analyses were performed to answer two main questions: (1) Do changes in EEG dynamics depend on the recording environment (lab vs. car)? (2) Is there a interaction between recording environment and feedback type (incorrect vs. correct)? A main effect of feedback type should be observed irrespective of the recording environment if EEG-recordings in a driving car reliably measure brain dynamics. A main effect of recording environment on P300 amplitudes, possibly reflecting decreased data quality due to in-vehicle artifact sources and movement of participants. Importantly, the absence of an interaction effect would indicate that the recording, analysis, and interpretation of EEG data in realistic driving scenarios is feasible for this particular task.

## 4.1. Data Analysis

Comprehensive analysis on both data sets recorded within the two recording environments were calculated. Differences in data characteristics in terms of trial amount for both recording environments were addressed. Tonic differences in power spectrum density ( $\mu$ V<sup>2</sup>/Hz) at midline electrode sites (Fz, Cz, Pz) were analyzed for the theta band (4 – 7 Hz) and alpha band (8 – 12 Hz). Power spectrum density estimates were calculated using Welch's method with windows of 256 points length, zero padded to 512 points and no overlap. Mean density values were assessed for both frequency bands by a 2 × 3 ANOVA with factors recording environment (lab, car) and electrode site (Fz, Cz, Pz). Event-related amplitude differences were assessed by a 2x2x3 mixed design ANOVA with the between factor recording environment (laboratory vs. car) and the within factors feedback type (correct vs. incorrect) and electrode site.

## 4.2. Results

#### 4.2.1. Data Characteristics

In total, significantly more trials were recorded in the lab environment (3,795 trials) than in the driving environment (2,185 trials),  $t_{(21.16)} = 14.11$ , p < 0.001. Furthermore, the proportion of trials rejected by automated cleaning was significantly higher,  $t_{(28)} = -2.84$ , p = 0.08, for epochs extracted from the driving study (28.89%) compared to the lab recordings (14.11%). Therefore, more trials were considered for analysis of the laboratory data (M = 217 trials per person, SD = 32.61) compared to the driving study data (M = 102 trials per person, SD = 23.30),  $t_{(28)} = 11.05$ , p < 0.001.



the upper plots are highlighted in gray in the ERP traces.

#### 4.2.2. Theta and Alpha Band Power

The analysis of the power in the theta frequency band revealed a significant main effect for the factor electrode site,  $F_{(2,56)} = 40.76, p < 0.001, \eta_p^2 = 0.59$  (Figure 5). Higher theta frequencies,  $t_{(29)} = 7.33, p < 0.001$ , were measured at frontal electrode site Fz ( $M = 1.76\mu V^2/Hz$ ,  $SD = 0.69\mu V^2/Hz$ ) compared to the central electrode site Cz ( $M = 0.98\mu V^2/Hz$ ,  $SD = 0.44\mu V^2/Hz$ ) and compared to parietal site Pz ( $M = 1.02\mu V^2/Hz$ ,  $SD = 0.45\mu V^2/Hz$ ),  $t_{(29)} = 6.90, p < 0.001$ . No difference,  $t_{(29)} = -0.45, p = 0.656$ , was found for theta frequencies at electrode site Cz compared to Pz. No main effect,  $F_{(1,28)} = 0.22, p < 0.642, \eta_p^2 = 0.01$ , or interaction effect,  $F_{(2,56)} = 2.35, p < 0.105, \eta_p^2 = 0.08$ , were found for the factor recording environment.

With respect to the power in the alpha frequency range, a significant main effect of electrode site was observed,  $F_{(2,56)} = 13.30, p < 0.001, \eta_p^2 = 0.32$ . This effect can be explained by reduced spectral power at central electrode sites ( $M = 0.64\mu V^2/Hz$ ,  $SD = 0.38\mu V^2/Hz$ ) as compared to frontal electrode sites ( $M = 1.02\mu V^2/Hz$ ,  $SD = 0.51\mu V^2/Hz$ ),  $t_{(29)} = 4.41, p < 0.001$ , and parietal electrode sites ( $M = 0.97\mu V^2/Hz$ ,  $SD = 0.54\mu V^2/Hz$ ),  $t_{(29)} = -4.12, p < 0.001$ . No difference,  $t_{(29)} = 0.62, p = 0.539$ , was found for alpha frequencies at electrode site Cz compared to Pz. Again, no main

effect,  $F_{(1,28)} = 1.90$ , p < 0.179,  $\eta_p^2 = 0.06$ , or interaction effect,  $F_{(2,56)} = 1.58$ , p < 0.217,  $\eta_p^2 = 0.05$ , were found for the factor recording environment.

#### 4.2.3. ERPs

The comparison analysis on both data sets revealed significant main effects for the factor feedback type,  $F_{(1,28)} = 75.60$ , p < 0.001,  $\eta_p^2 = 0.73$ , and electrode site,  $F_{(1.36,38.02)} = 17.40$ , p < 0.001,  $\eta_p^2 = 0.38$ . Mean P300 amplitudes elicited by incorrect feedback ( $M = 2.13\mu$ V,  $SD = 1.46\mu$ V) were more pronounced compared to correct feedback ( $M = -0.34\mu$ V,  $SD = 1.09\mu$ V). Activity at Fz ( $M = -0.59\mu$ V,  $SD = 2.16\mu$ V) was lower than activity recorded at Cz ( $M = 1.04\mu$ V,  $SD = 1.64\mu$ V),  $t_{(29)} = -4.76$ , p < 0.001, and Pz ( $M = 2.24\mu$ V,  $SD = 1.73\mu$ V),  $t_{(29)} = -4.53$ , p < 0.001 and lower at Cz compared to Pz,  $t_{(29)} = -2.60$ , p = 0.015.

The main effects were qualified by a significant interaction of the factors feedback type and electrode site,  $F_{(2,56)} = 7.84$ , p = 0.001,  $\eta_p^2 = 0.22$ , revealing highest P300 amplitudes for correct feedback at Pz ( $M = 0.62\mu$ V,  $SD = 1.60\mu$ V) compared to Fz ( $M = -1.09\mu$ V,  $SD = 2.03\mu$ V),  $t_{(29)} = -3.22$ , p = 0.003. Incorrect feedback elicited lower P300 amplitudes at Fz ( $M = -0.07\mu$ V,  $SD = 2.84\mu$ V) compared to Cz ( $M = 2.71\mu$ V,  $SD = 2.84\mu$ V),  $t_{(29)} = -5.37$ , p < 0.001, and Pz



 $(M = 3.95\mu$ V,  $SD = 2.23\mu$ V),  $t_{(29)} = -4.72$ , p < 0.001. Amplitudes for incorrect feedback were larger than for correct feedback at all three electrode sites, Fz:  $t_{(29)} = -2.81$ , p = 0.009, Cz:  $t_{(29)} = -5.05$ , p < 0.001, Pz:  $t_{(29)} = -11.30$ , p < 0.001.

A marginal significant main effect of recording environment,  $F_{(1,28)} = 3.42$ , p = 0.075,  $\eta_p^2 = 0.11$ , with higher amplitudes in the laboratory study ( $M = 1.14\mu$ V,  $SD = 0.87\mu$ V) compared to the driving study ( $M = 0.65\mu$ V,  $SD = 1.14\mu$ V) was specified by an again marginal interaction effect of recording environment x electrode site,  $F_{(2,56)} = 3.09$ , p = 0.053,  $\eta_p^2 = 0.10$ . A tendency toward larger amplitudes at electrode site Pz was found for the laboratory study ( $M = 2.97\mu$ V,  $SD = 0.91\mu$ V) compared to the driving study ( $M = 1.50\mu$ V,  $SD = 2.05\mu$ V),  $t_{(19.30)} = 2.55$ , p =0.019. No significant effects were found for the interaction of feedback type × recording environment,  $F_{(1,28)} = 1.26$ , p =0.271,  $\eta_p^2 = 0.04$ , and feedback type × electrode site × recording environment,  $F_{(2,56)} = 1.94$ , p = 0.153,  $\eta_p^2 = 0.07$ .

#### 4.3. General Discussion

Two studies were conducted to establish an experimental protocol for systematically comparing the neural responses elicited by unexpected erroneous events within a realistic driving setting. In the first study we tested our experimental manipulation successfully by provoking the well-known P300 deflection for the processing of infrequent but task-relevant auditory events (Sutton et al., 1965; Katayama and Polich, 1996). In a second study, the same test was carried out in a real driving scenario, replicating the P300 response observed in the first study.

While the results demonstrate that it is feasible to investigate the neural dynamics underlying incorrect feedback processing in both scenarios, general differences in data characteristics had to be addressed for a more specific comparison. Despite clear visual similarities in mean ERP traces from both acquisitions (shown in **Figure 6**), higher variance was observed in the data recorded in the car. As a real-life driving scenario is an inherent source of technical artifacts and active behavior, differences in signal quality are not unexpected. This was confirmed by a significantly higher number of trials subject to automated artifact rejection due to amplitudes that exceeded a criterion of  $+80 \ \mu V$  or deviated clearly from the mean calculated probability distribution. Moreover, the more complex and time consuming preparation and acquisitions sessions in the driving setup led to generally shorter recording times. These two factors accounted for a significant lower number of trials for the in-car recordings. Furthermore, the introduction of a perceptual demanding additional driving task could have influenced P300 amplitudes (Wickens et al., 1983) and increased the variance in the driving condition. However, clear P300 deflections for incorrect feedback events were observed, as before in the laboratory assessment. To allow a more direct comparison, further analyses in the time and frequency domain were computed with both data sets. While a clear P300 component associated with the processing of infrequent and task-relevant stimuli was successfully replicated under realistic driving conditions, no significant main or interaction effect of the factor recording environment on P300 amplitudes was revealed by the analysis. Furthermore, the general impact of data quality in the different recording environments on the P300 deflection was addressed by analysis in the frequency domain. Tonic power spectra in both recordings were comparable and again, no significant effect of the factor recording environment was found.

Following the argumentation on a reciprocal relationship between task difficulty of the primary task and P300 amplitudes in the secondary task (e.g., Sirevaag et al., 1989) we suggest that the task load introduced by the driving task was not sufficient to produce a significant effect. Thus, keeping track and speed on a straight and blocked course appears to demand little attentional resources. However, marginal significant differences between overall amplitudes in both recording environments and a marginal significant interaction effect of recording environment and electrode site indicate potential primary task



resource demands. Mean amplitudes tended to be smaller at parietal electrode sites while driving compared to the laboratory assessment. This tendency might reflect the reallocation of resources that were needed to accomplish the driving task. We expect more complex driving tasks to result in a pronounced reduction of P300 amplitudes. In addition, future studies might also consider a within-subject design to control for possible group differences that could have caused the obtained marginal effects of the different recording environments on parietal activity.

In sum, our approach showed that P300 amplitudes elicited by unexpected and erroneous events can be assessed during the performance of an unhindered driving task. Once more, the feasibility of EEG measurements beyond more or less restricted standard laboratory settings with new applicationoriented approaches was demonstrated in this study (e.g., Gwin et al., 2010; Debener et al., 2012; Jungnickel and Gramann, 2016). Based on these results, more complex dual-task paradigms with varied difficulty levels in either the primary driving task or in the secondary task can be addressed. This will be of importance for further research in autonomous driving and for the development of driving assistance by providing insights into the driver's processing of incoming information while interacting with the car and the surrounding environment. Thus, systematic analysis on variations in different stages of information processing could be used for more direct driver state assessments and the design of adaptive assistance.

## **5. CONCLUSION**

With two studies we were first able to replicate previous laboratory- based work on P300 amplitudes and then to confirm a high level of ecological validity of our results in a realistic driving task setting. Our findings provide strong evidence that complex cognitive functions like context and response updating processes can be examined in a highly artifact prone driving environment. The processing of infrequent and incorrect auditory feedback events was reflected by clear P300 deflections with slightly different topographical distributions in both recordings. Despite differences in data quality and variance, amplitudes and tonic EEG power spectra from both studies were comparable and not significantly affected by the factor recording environment. The possibilities to provide direct insights into brain dynamics of humans participating in a real world driving task provides compelling arguments for further investigation in realistic task settings with more complex manipulation or on less robust potentials. A gradual transfer

## REFERENCES

- Banaei, M., Hatami, J., Yazdanfar, A., and Gramann, K. (2017). Walking through architectural spaces: the impact of interior forms on human brain dynamics. *Front. Hum. Neurosci.* 11:477. doi: 10.3389/fnhum.2017.00477
- Brouwer, A.-M., Snelting, A., Jaswa, M., Flascher, O., Krol, L., and Zander, T. (2017). "Physiological effects of adaptive cruise control behaviour in real driving," in *Proceedings of the 2017 ACM Workshop on An Application-oriented Approach to BCI out of the Laboratory* (Limassol: ACM), 15–19.
- Chatrian, G., Lettich, E., and Nelson, P. (1985). Ten percent electrode system for topographic studies of spontaneous and evoked EEG activities. Am. J. EEG Technol. 25, 83–92. doi: 10.1080/00029238.1985.11080163
- Courchesne, E. (1978). Changes in P3 waves with event repetition: longterm effects on scalp distribution and amplitude. *Electroencephalogr. Clin. Neurophysiol.* 45, 754–766. doi: 10.1016/0013-4694(78)90143-8
- Dahlbäck, N., Jönsson, A., and Ahrenberg, L. (1993). Wizard of Oz studies why and how. *Knowl. Based Syst.* 6, 258–266. doi: 10.1016/0950-7051(93)90017-N
- de Bruijn, E. R., Schubotz, R. I., and Ullsperger, M. (2007). An event-related potential study on the observation of erroneous everyday actions. *Cogn. Affect. Behav. Neurosci.* 7, 278–285. doi: 10.3758/CABN.7.4.278
- Debener, S., Minow, F., Emkes, R., Gandras, K., and Vos, M. (2012). How about taking a low-cost, small, and wireless EEG for a walk? *Psychophysiology* 49, 1617–1621. doi: 10.1111/j.1469-8986.2012.01471.x
- Delorme, A., and Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *J. Neurosci. Methods* 134, 9–21. doi: 10.1016/j.jneumeth.2003.10.009
- Donchin, E., and Coles, M. G. (1988). Is the P300 component a manifestation of context updating? *Behav. Brain Sci.* 11, 357–374. doi: 10.1017/S0140525X00058027
- Donchin, E., Ritter, W., and McCallum, W. C. (1978). "Cognitive psychophysiology: the endogenous components of the ERP," in *Event-Related Brain Potentials in Man*, ed E. Callaway (New York, NY: Academic Press), 349–411.
- Engel, A. K., Maye, A., Kurthen, M., and König, P. (2013). Where's the action? The pragmatic turn in cognitive science. *Trends Cogn. Sci.* 17, 202–209. doi: 10.1016/j.tics.2013.03.006
- Fabiani, M., Gratton, G., Karis, D., and Donchin, E. (1987). "Definition, identification, and reliability of measurement of the P300 component of the event-related brain potential," in *Advances in psychophysiology, Vol.2*, eds P. K. Ackles, J. R. Jennings, and M. G. H. Coles (Greenwich, CT: JAI Press), 1–78.

of the extensive knowledge gathered from laboratory ERP reports into ecological task settings could prospectively result in complex findings about brain dynamics of actively behaving humans.

## **AUTHOR CONTRIBUTIONS**

JP carried out the experiment. JP and KG contributed to the analysis and interpretation of the data. JP and KG wrote the manuscript.

## ACKNOWLEDGMENTS

The present project was embedded in the Research Training Group *prometei* - *Prospective Design of Human-Technology Interaction* founded by the German Research Foundation (DFG) and the data recordings were parts of a dissertation project (Protzak, 2015). A preprint version of this article (Protzak and Gramann, 2018) was posted on the bioRxiv.org pre-print server.

- Friedman, D., Cycowicz, Y. M., and Gaeta, H. (2001). The novelty P3: an eventrelated brain potential (ERP) sign of the brain's evaluation of novelty. *Neurosci. Biobehav. Rev.* 25, 355–373. doi: 10.1016/S0149-7634(01)00019-7
- Gramann, K. (2013). Embodiment of spatial reference frames and individual differences in reference frame proclivity. Spat. Cogn. Comput. 13, 1–25. doi: 10.1080/13875868.2011.589038
- Gramann, K., Ferris, D. P., Gwin, J., and Makeig, S. (2014). Imaging natural cognition in action. *Int. J. Psychophysiol.* 91, 22–29. doi: 10.1016/j.ijpsycho.2013.09.003
- Gramann, K., Gwin, J. T., Ferris, D. P., Oie, K., Jung, T. P., Lin, C. T., et al. (2011). Cognition in action: imaging brain/body dynamics in mobile humans. *Rev. Neurosci.* 22, 593–608. doi: 10.1515/RNS.2011.047
- Gwin, J. T., Gramann, K., Makeig, S., and Ferris, D. P. (2010). Removal of movement artifact from high-density EEG recorded during walking and running. J. Neurophysiol. 103, 3526–3534. doi: 10.1152/jn.00105.2010
- Haufe, S., Kim, J. W., Kim, I. H., Sonnleitner, A., Schrauf, M., Curio, G., et al. (2014). Electrophysiology-based detection of emergency braking intention in real-world driving. *J. Neural Eng.* 11:056011. doi: 10.1088/1741-2560/11/5/056011
- Haufe, S., Treder, M. S., Gugler, M. F., Sagebaum, M., Curio, G., and Blankertz, B. (2011). EEG potentials predict upcoming emergency brakings during simulated driving. *J. Neural Eng.* 8:056001. doi: 10.1088/1741-2560/8/5/056001
- Horst, R. L., Johnson, R., and Donchin, E. (1980). Event-related brain potentials and subjective probability in a learning task. *Mem. Cogn.* 8, 476–488. doi: 10.3758/BF03211144
- Isreal, J. B., Chesney, G. L., Wickens, C. D., and Donchin, E. (1980a). P300 and tracking difficulty: evidence for multiple resources in dual-task performance. *Psychophysiology* 17, 259–273. doi: 10.1111/j.1469-8986.1980.tb00146.x
- Isreal, J. B., Wickens, C. D., Chesney, G. L., and Donchin, E. (1980b). The eventrelated brain potential as an index of display-monitoring workload. *Hum. Factors* 22, 211–224. doi: 10.1177/001872088002200210
- Johnson, R. (1993). On the neural generators of the P300 component of the event-related potential. *Psychophysiology* 30, 90–97. doi: 10.1111/j.1469-8986.1993.tb03208.x
- Jungnickel, E., and Gramann, K. (2016). Mobile brain/body imaging (MoBI) of physical interaction with dynamically moving objects. *Front. Hum. Neurosci.* 10:306. doi: 10.3389/fnhum.2016.00306
- Katayama, J., and Polich, J. (1996). P300, probability, and the threetone paradigm. *Electroencephalogr. Clin. Neurophysiol.* 100, 555–562. doi: 10.1016/S0168-5597(96)95171-0

- Kecklund, G., and Akerstedt, T. (1993). Sleepiness in long distance truck driving: an ambulatory EEG study of night driving. *Ergonomics* 36, 1007–1017. doi: 10.1080/00140139308967973
- Kohlmorgen, J., Dornhege, G., Braun, M., Blankertz, B., Müller, K.-R., Curio, G., et al. (2007). "Improving human performance in a real operating environment through real-time mental workload detection," in *Toward Brain-Computer Interfacing*, eds G. Dornhege, J. d. R. Millan, T. Hinterberger, D. McFarland, and K.-R. Mueller (Cambridge, MA: MIT Press), 409–422.
- Krol, L., Zander, T., Jaswa, M., Flascher, O., Snelting, A., and Brouwer, A. (2017). "Online-capable cleaning of highly artefactual EEG data recorded during real driving," in *Proceedings of the 7th Graz Brain-Computer Interface Conference* 2017 (Graz), 254–259.
- Makeig, S., Gramann, K., Jung, T.-P., Sejnowski, T. J., and Poizner, H. (2009). Linking brain, mind and behavior. *Int. J. Psychophysiol.* 73, 95–100. doi: 10.1016/j.ijpsycho.2008.11.008
- Oldfield, R. C. (1971). The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia* 9, 97–113. doi: 10.1016/0028-3932(71)90067-4
- Palmer, J. A., Makeig, S., Kreutz-Delgado, K., and Rao, B. D. (2008). "Newton method for the ICA mixture model," in *IEEE International Conference* on Acoustics, Speech and Signal Processing, 2008 (Las Vegas, NV: IEEE), 1805–1808.
- Papadelis, C., Chen, Z., Kourtidou-Papadeli, C., Bamidis, P. D., Chouvarda, I., Bekiaris, E., et al. (2007). Monitoring sleepiness with on-board electrophysiological recordings for preventing sleep-deprived traffic accidents. *Clin. Neurophysiol.* 118, 1906–1922. doi: 10.1016/j.clinph.2007.04.031
- Picton, T. W. (1992). The P300 wave of the human event-related potential. J. Clin. Neurophysiol. 9, 456–456. doi: 10.1097/00004691-199210000-00002
- Polich, J. (2007). Updating P300: an integrative theory of P3a and P3b. Clin. Neurophysiol. 118, 2128–2148. doi: 10.1016/j.clinph.2007.04.019
- Polich, J., Eischen, S. E., and Collins, G. E. (1994). P300 from a single auditory stimulus. *Electroencephalogr. Clin. Neurophysiol.* 92, 253–261. doi: 10.1016/0168-5597(94)90068-X
- Protzak, J. (2015). *EEG*-Korrelate der Fehlerverarbeitung während der Fahraufgabe [Dissertation], Technische Universität, Berlin.
- Protzak, J., and Gramann, K. (2018). Investigating established EEG parameter during real-world driving. *BioRxiv* 275396[Preprint]. doi: 10.1101/275396
- Schmidt, E. A., Schrauf, M., Simon, M., Fritzsche, M., Buchner, A., and Kincses, W. E. (2009). Drivers' misjudgement of vigilance state during prolonged monotonous daytime driving. *Accid Anal. Prev.* 41, 1087–1093. doi: 10.1016/j.aap.2009.06.007

- Segalowitz, S. J., Wintink, A. J., and Cudmore, L. J. (2001). P3 topographical change with task familiarization and task complexity. *Cogn. Brain Res.* 12, 451–457. doi: 10.1016/S0926-6410(01)00082-9
- Simon, M., Schmidt, E. A., Kincses, W. E., Fritzsche, M., Bruns, A., Aufmuth, C., et al. (2011). EEG alpha spindle measures as indicators of driver fatigue under real traffic conditions. *Clin. Neurophysiol.* 122, 1168–1178. doi: 10.1016/j.clinph.2010.10.044
- Sirevaag, E. J., Kramer, A. F., Coles, M. G., and Donchin, E. (1989). Resource reciprocity: an event-related brain potentials analysis. *Acta Psychol.* 70, 77–97. doi: 10.1016/0001-6918(89)90061-9
- Sonnleitner, A., Treder, M. S., Simon, M., Willmann, S., Ewald, A., Buchner, A., et al. (2014). EEG alpha spindles and prolonged brake reaction times during auditory distraction in an on-road driving study. Accid Anal. Prev. 62, 110–118. doi: 10.1016/j.aap.2013. 08.026
- Sutton, S., Braren, M., Zubin, J., and John, E. (1965). Evokedpotential correlates of stimulus uncertainty. *Science* 150, 1187–1188. doi: 10.1126/science.150.3700.1187
- Verleger, R., Jaśkowski, P., and Wascher, E. (2005). Evidence for an integrative role of P3b in linking reaction to perception. J. Psychophysiol. 19, 165–181. doi: 10.1027/0269-8803.19.3.165
- Wickens, C., Kramer, A., Vanasse, L., and Donchin, E. (1983). Performance of concurrent tasks: a psychophysiological analysis of the reciprocity of information-processing resources. *Science* 221, 1080–1082. doi: 10.1126/science.6879207
- Zhang, H., Chavarriaga, R., Khaliliardali, Z., Gheorghe, L., Iturrate, I., and Millán, J. D. R. (2015). EEG-based decoding of error-related brain activity in a realworld driving task. *J. Neural Eng.* 12:066028. doi: 10.1088/1741-2560/12/6/ 066028

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2018 Protzak and Gramann. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.