

Correlation of Subjective Expectation and P300 Amplitude during a Game of Matching Pennies^{*}

Arthur Flexer^{a,b,*},¹, Scott Makeig^a

^a*Swartz Center for Computational Neuroscience, Institute for Neural Computation, University of California San Diego 0961, La Jolla, CA 92093-0961, USA²*

^b*The Austrian Research Institute for Artificial Intelligence, Freyung 6/6, A-1010 Vienna, Austria³*

Abstract

We report a study on two-person game playing involving simultaneous EEG recording from both subjects. Independent Component Analysis is used for identifying activities of individual cortical EEG sources. Activity of a midline fronto-central component identified in four of five subjects was correlated with a measure of subjective expectation. This component accounts for the P300 waveform whose amplitude varies depending on the context of the gaming situation.

Key words: Independent Component Analysis, EEG, Neuroeconomics

* Thanks to Julie Onton and Andrey Vankov for helping with the experiments.

* Corresponding author. Tel.: +43 - 1 - 533 6112, Fax.: +43 - 1 - 533 6112 - 77.

Email addresses: arthur@oefai.at (Arthur Flexer), scott@sccn.ucsd.edu (Scott Makeig).

¹ Arthur Flexer is supported by an Erwin Schrödinger Fellowship provided by the Austrian Science Fund (FWF), project J2221-N04.

² The Swartz Center receives support from the Swartz Foundation, East Setauket NY.

³ The Austrian Research Institute for Artificial Intelligence is supported by the Austrian Federal Ministry of Education, Science and Culture and the Austrian Federal Ministry for Transport, Innovation and Technology.

1 Introduction

This report is about first results in a neuroscientific study of two-person game playing. A number of recent studies on game playing monkeys using implanted electrodes have produced new insights into the function of the brain. These types of studies have successfully linked single cell firing rates to mathematical behavioral models, thereby producing a new field of research sometimes termed neuroeconomics (see [Glimcher 2003] for a comprehensive introduction). Our experiment seeks to clarify whether a similar study can find anticipatory evaluative signals in scalp EEG signals recorded non-invasively from humans.

In our EEG experiment two subjects are tested at once, competing against one another in a simple "Same/Different" game formally akin to the classical game of "Matching Pennies" (see e.g. [Camerer 2004]). We use Independent Component Analysis (ICA) (see [Comon 1994], [Bell & Sejnowski 1995] and [Makeig et al. 1996]) to decompose EEG recorded during the game into non-brain artifacts and spontaneous EEG activity. Analysing the activity of the independent components instead of the raw EEG channel time series gives a much clearer and more artifact-free picture of the activities of individual cortical sources [Makeig et al. 2002]. During the games, the largest non-artifact component in all but one subject had a midline fronto-central localization and gave rise to a late positive component after auditory feedback. We show that a measure of subjective expectation ("negative surprise") depending on the recent history of wins and losses in the game is correlated to the amplitude of this component, which accounts for a portion of the well known P300 waveform [Makeig et al. in press]. This connects our research on game playing to previous results on the effect of stimulus sequence on P300 waveforms [Squires et al. 1976].

2 Methods

2.1 *Game*

The recording of the human electroencephalogram (EEG) is a non-invasive method of recording electric brain potentials from the human scalp via a set of electrodes. In our EEG experiment two subjects were tested at once, competing against one another in a simple "Same/Different" game. During these experiments, each subject wore 128 EEG electrodes positioned evenly across the head, plus four sub-ocular electrodes recording their eye movements (EOG). Simultaneous recording of their EEG and EOG was acquired using a custom two-subject Biosemi EEG system. We recorded data from three pairs of sub-

jects (male, right handed, age = 29.6 ± 9.29 *sd*) who did not know each other prior to the experiment. Data were recorded with a sampling rate of 256 *Hz*, FIR-filtered with a bandpass from 1 to 50*Hz* and then down-sampled to 128 *Hz*.

Both subjects were seated facing one another in the experiment room. Subjects were asked to avoid communicating to each other during the experimental session, including during breaks. Between the subjects, two LCD screens, placed below face level, were used to deliver visual feedback. Each subject held a response button in their dominant hand. The two LCD screens shielded each participant's view of the other participant's hand and button, without blocking the view of the other subject's face. Subjects were instructed to look at their competitor's face when actually playing. Each subject wore ear insert headphones that delivered individual performance feedback during the game.

Participants competed against one another in a simple "Same/Different" game formally akin to the classical game of "Matching Pennies". Matching Pennies is a zero-sum game with two players (see e.g. [Camerer 2004]). In each play, each player shows one side of a coin, either heads or tails, to the other. If both players' coins show heads or both tails, then player One wins, otherwise player Two wins. Matching Pennies is one in a family of games in which one person benefits from a "match" whereas the other benefits from a "mismatch". Rational human beings, when confronted with this strategic conflict reach a behavioral equilibrium at which the average subjective desirability of the two actions comes to equivalence (see e.g. [Nash 1950]).

Participants were first taught the game including a short practice period. They were told whether they had been assigned to play the "Same" or the "Different" strategy. During the game, the "Same" player was rewarded with points and positive auditory feedback for having the response button in the same position as that of their competitor (up = up or down = down). The "Different" player was rewarded for having it in the different position from that of their competitor (up \neq down or down \neq up). At the end of the game playing session a bonus of \$25 was divided between both players according to the ratio of points earned during the game. If one of the players ended the game with zero points or a negative score, the other one received the whole bonus. Any bonus was in addition to the standard hourly rate. The pseudo-code like representation in Tab. 1, plus the next paragraphs more formally describe the game.

The game session consisted of four BLOCKs of ten BOUTs. Each BOUT consisted of twenty PLAYS. Each BOUT started with the instruction "READY" displayed on both LCD screens for two seconds (`display(READY,2sec)`). This was followed automatically by the display of the instruction "PLAY" (`display(PLAY,until BOUT end)`) which stayed on the screens for the rest

Table 1

Outline of “Same/Different” game.

```

for BLOCK = 1 : 4
  for BOUT = 1 : 10
    display(READY,2sec)
    display(PLAY,until BOUT end)
    for PLAY = 1 : 20
      wait(1-5sec)
      evaluate_and_feedback
    end
    display(RESULT-BOUT,until button presses)
    display(RESULT-TOTAL,until button presses)
  end
  if not (BLOCK == 4)
    display(RELAX,until instructor restart)
  end
end
display(RESULT-FINAL)

```

of the BOUT.

Players then covertly pressed and released their response button at self-chosen intervals to maximize the chance of holding it either in the same or different position than the other player. At irregular, 1-5 sec intervals chosen at random by the computer running the game (`wait(1-5sec)`), the computer tallied the relative position of the two participants’ buttons, computed points won or lost by each subject, and rewarded one participant with positive auditory feedback (“beep”⁴) and the other with concurrent negative feedback (“buzz”⁵) (`evaluate-and-feedback`). To minimize rapid button pressing by either subject, points awarded for each winning PLAY, and lost for each losing PLAY, were proportional to the length of time the subject had held the button in the same position (to a maximum of 5 sec). To minimize the chance of any participant ending the session with a negative point total, points lost in a PLAY increased with hold time at a slower rate than points won. Larger point gains are indicated to subjects with higher-pitched (“beep”) feedback; larger point losses were indicated by lower-pitched (“buzz”) feedback⁶.

After twenty PLAYS, the BOUT was completed. The two LCD screens indicated the total number of points won/lost during the last BOUT (`display(RESULT-BOUT,until button presses)`), and then the scoring in the session so far (`display(RESULT-TOTAL,until button presses)`). The

⁴ Sinus wave, .2 sec duration, 1300 *Hz* frequency, 10% on- and off-ramp.

⁵ Sawtooth wave, .2 sec duration, 500 *Hz* frequency, 10% on- and off-ramp.

⁶ Adding or subtracting 50 or 100 *Hz* to the basic “beep” or “buzz” sounds.

last display also showed each of the players the share of the bonus (in US dollars) they would earn if the game ended right then. The next BOUT began after both players pressed their response button at least once.

After twenty BOUT periods a game BLOCK was completed. Between BLOCKS, subjects were allowed to relax (if not (BLOCK == 4)) by stretching, closing their eyes, etc. Players were allowed to talk to the instructor but not to each other. This was indicated by displaying the instruction “RELAX” on the LCD screens (display(RELAX,until instructor restart)). The game resumed after both players told the instructor they were ready to begin again. After four game BLOCKS, the LCD screen displayed final scores and shares of the bonus to both players (display(RESULT-FINAL)).

One complete “Same/Different” session required about $4 \times 10 \times 20$ PLAYS $\times 3sec = 40min$ not including break times between blocks and bouts. Therefore the game session lasted less than an hour.

2.2 Preprocessing

Before any further processing of the data, channels which were noisy for the majority of the recording period (e.g. due to drift or bad contact) and segments containing severe artifacts (disturbances across many channels) were deleted. One subject had to be excluded from further analysis because of problems with the ear insert headphones during the game. This left five subjects for the analysis.

2.3 Analysis

Rather than analysing the raw EEG we used Independent Component Analysis (ICA) to decompose the EEG recorded during the game into artifacts and spontaneous EEG activity. This allows us to get a clearer and more artifact-free picture of the activities of individual cortical sources [Makeig et al. 2004].

Independent Component Analysis (ICA) [Comon 1994] is one of a group of algorithms that attempt to achieve blind separation of sources [Jutten & Herault 1991]. To estimate the original sources from an observed mixture, while knowing little about the mixing process and making only few assumptions about it and about the sources, is called blind separation of sources. ICA allows recovery of N independent source signals $s = \{s_1(t), s_2(t), \dots, s_N(t)\}$ from N linear mixtures, $x = \{x_1(t), x_2(t), \dots, x_N(t)\}$, modeled as the result of multiplying the matrix of source activity waveforms, s , by an unknown square matrix A (i.e. $x = As$). The task is to recover a

version, u , of the original sources s , save for scaling and ordering. To this purpose, it is necessary to find a square matrix W specifying filters that linearly invert the mixing process (i.e. $u = Wx$).

By the central limit theorem a linear mixture of independent random variables is necessarily more Gaussian than the original variables. Therefore maximizing the nongaussianity achieves the unmixing of the recorded signals x . This implies (i) that it is enough to assume that the source signals $s_i(t)$ are *statistically independent* at each time step t , though their mixtures $x_i(t)$ are not; (ii) that in ICA we must restrict ourselves to at most one Gaussian source signal. Since there exist numerous ways to measure nongaussianity (e.g. kurtosis, negentropy, etc.) and different approaches towards information maximization, ICA researchers have developed a family of algorithms for solving the blind source separation problem (see e.g. [Lee 1998], [Hyvaerinen 1999] and [Roberts & Everson 2001] for an introduction and overview).

We used the "infomax" neural network algorithm [Bell & Sejnowski 1995] for ICA as implemented by [Makeig et al. 1997]⁷. This approach uses the fact that maximizing the joint entropy, $H(y)$, of the output of a neural processor minimizes the mutual information among the output components, $y_I = g(u_i)$, where $g(u_i)$ is an inverted bounded nonlinearity and $u = Wx$.

ICA has already been used successfully for blind source separation of EEG data. Applications of ICA to EEG include artifact detection and removal (see [Makeig et al. 1996], [Jung et al. 1998] and [Jung et al. 2000]) as well as analysis of event-related response averages (see [Makeig et al. 1997], [Makeig et al. 1999a] and [Makeig et al. 1999b]). Application of ICA to continuous EEG is more recent (see [Jung et al. 1999], [Jung et al. 2001] and [Makeig et al. 2002]). In continuous EEG analysis, the rows of the input matrix x are EEG and EOG signals recorded at different electrodes and the columns are measurements at different time points. ICA finds an unmixing matrix W which linearly decomposes the multichannel data into a sum of maximally temporally independent and spatially fixed components $u = Wx$. The rows of the output matrix u are courses of activation of the independent components (ICs). These components account for artifacts, stimulus and response locked events and spontaneous EEG activity. The columns of the inverse matrix W^{-1} give the relative projection strengths of the respective ICs at each of the scalp sensors. These scalp maps of projection strengths provide evidence for the components' physiological origin (e.g. ocular activity projects mainly to frontal sites). Selected components can be projected back onto the scalp using the relation $x_0 = W^{-1}u_0$, where u_0 is the matrix u with irrelevant components set to zero. Thereby brain signals accounted for by the selected

⁷ All ICA related computations were done with the MATLAB toolbox EEGLAB [Delorme & Makeig 2004].

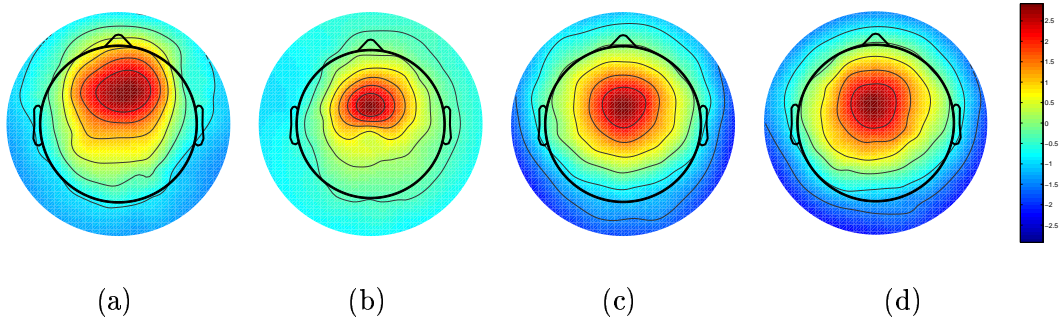


Fig. 1. Topographical plots of largest non-artifact independent components for subject1 (a), subject2 (b), subject3 (c), subject4 (d). Given in z-scores, red being positive and blue negative.

components can be obtained in their true polarity and amplitudes.

3 Results

We computed infomax ICA for each of the five subjects separately. The respective input matrices x consisted of all EEG and EOG channels from one subject remaining after manual artifact inspection (see end of Sec. 2.1). We used principal component analysis to reduce the dimensionality of the input channels to 90. We found this to be advantageous for numerical reasons. ICA outputs the IC activations u and the square matrix W specifying the filters that invert the mixing process ($u = Wx$). We inspected topographical plots of all IC scalp maps (the columns of the inverse matrix W^{-1}) for all subjects after sorting the components in descending order of their mean projected variance (using $x_0 = W^{-1}u_0$). For four of the five subjects, we found a component with midline fronto-central localization to be a principal non-artifact component (artifact components account for eye movements, muscle noise etc.). For three subjects this component was the largest non-artifact component, for one the second largest. Topographical plots of these four components are depicted in Fig. 1. Since ICA estimates u and W only up to an arbitrary scaling factor, topographical plots are made after transformation to z-scores to allow for comparability across subjects.

In a next step we epoched the IC activations u of each of the four subjects around the auditory feedback events. At the end of each PLAY there was an `evaluate-and-feedback-event` consisting of an auditory feedback in the form of a “beep” (positive for a win) or a “buzz” (negative for a loss). Within each experiment there were 800 such stimuli, the number being decreased by artifact rejection later on. A “feedback-locked epoch” consisted of the activation of an IC one second prior to two seconds after the feedback stimuli. Average

activations for the four ICs shown in Fig. 1 time-locked to the feedback event, are given in Fig. 2.

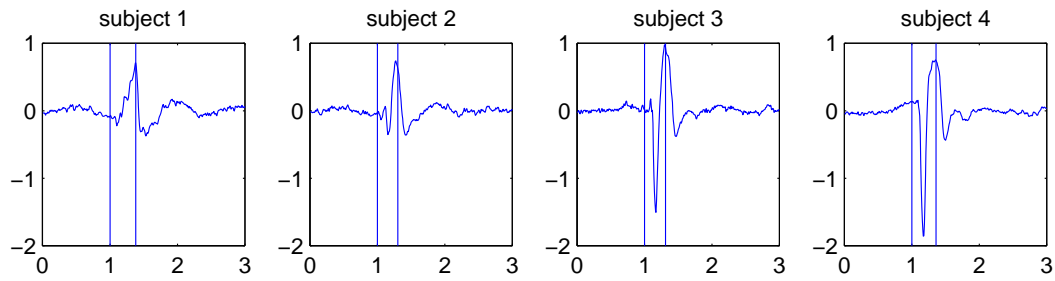


Fig. 2. Average IC activations, x-axis are seconds, y-axis amplitude of activations, first vertical line in each plot is stimulus onset, second is estimated peak of P300 activity.

The most prominent feature of all four averages seems to be a positive going waveform starting about 300 *ms* after stimulus onset. We estimated the peak of this activity by finding the maximum of each average between 300 and 400

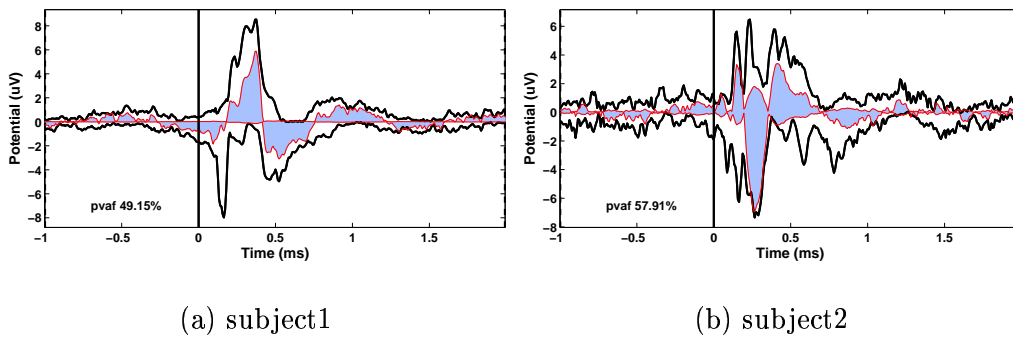


Fig. 3. Envelopes of averaged EEG time-locked to the the feedback event (outer lines), plus envelope of the backprojections of the selected components (shaded area) for subjects 1 and 2.

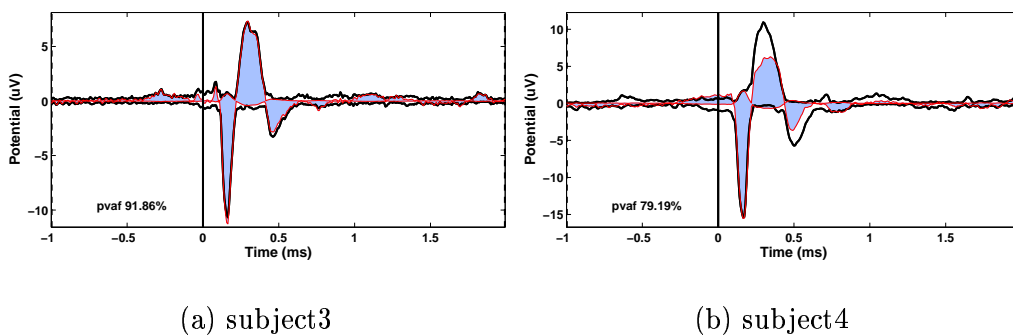


Fig. 4. Envelopes of averaged EEG time-locked to the the feedback event (outer lines), plus envelope of the backprojections of the selected components (shaded area) for subjects 3 and 4.

ms after stimulus onset. The second vertical line in each plot indicates this maximum. The average latency of these peaks is $339.8 ms \pm 37.5 sd$ after stimulus onset. This positive going waveform accounted for a portion of the well known P300 waveform (see e.g. [Sutton et al. 1965]). In Figs. 3 and 4 the envelope of the averaged EEG time-locked to the the feedback event, plus the envelope of the backprojections of the selected components are plotted for all four subjects. The positive going waveform before $500 ms$ is the P300, the shaded area shows to what extent an component accounted for it. The same spatial components also explained some or all of the N1 auditory response. The differences in spatial complexity of the auditory evoked responses in the four subjects are marked, but the significance of this difference is unknown.

Each of the feedback events is equivalent to a win or a loss for the subjects. Our next goal was to try to find a relation between the context of wins and losses to the IC activations. We used the following estimate of the expectancy of future wins and losses:

$$\hat{E}_{(t)} = \sum_{c=1}^C \alpha_{(t-c)} E_{(t-c)} \quad (1)$$

where t is a time index, C is the size of the context, α a decaying contribution of prior feedback stimuli ($\sum \alpha = 1$) and $E_{(i)} = 1$ if there was a loss at time i and $E_{(i)} = 0$ if there was a win at time i . $\hat{E}_{(t)}$ equals 1 when a loss is expected, 0 when a win is expected, and inbetween when the expectancy is less clear. Contrasting the expectancy $\hat{E}_{(t)}$ with the actual PLAY feedback $E_{(t)}$ gives an index of “negative surprise” (NS):

$$NS = E_{(t)} - \hat{E}_{(t)} \quad (2)$$

The index NS is maximal and equal 1 if all prior feedback stimuli are wins and the present feedback stimuli is a loss (e.g. for $C = 2$: win-win-loss). It is minimal and equal -1 for the reverse situation (e.g. for $C = 2$: loss-loss-win) and equal 0 in case the expectancy is fulfilled (e.g. for $C = 2$: loss-loss-loss or win-win-win).

For each of the four subjects we computed averages across the activations in single trials of their respective fronto-central components separately for different contexts. This gave, per subject, two averages for a context of size $C = 0$ (win or loss at the present time, W or L), four averages for a context of size $C = 1$ (WW,LW,WL,LL) and eight averages for a context of size $C = 2$ (WWW,LWW,WLW,LLW,WWL,LWL,WLL,LLL). For context $C = 1$ we set $\alpha_{(t-1)} = 1$ and for context $C = 2$ we set $\alpha_{(t-1)} = \frac{2}{3}$ and $\alpha_{(t-2)} = \frac{1}{3}$. For each of the averages we computed an estimate of P300 amplitude, $P300amp$, by

taking the mean of the average IC activation from 50 *ms* before to 50 *ms* after its peak (see above for how the peak was found). We subtracted the mean of a one second pre-stimulus baseline from this P300 amplitude estimate. Since ICA estimates the IC activations u only up to an arbitrary scaling factor, we have to normalize the $P300amp$ estimates to allow for comparison across subjects. We did this by computing a $P300amp$ estimate from the average across all feedback epochs from a subject (no matter whether it was win or loss or regardless of context, see Fig. 2), setting this to be 100% and expressing all other $P300amp$ as percentages relative to this overall estimate. In Fig. 5 (left) we plotted wins and losses versus $P300amp$ for context $C = 0$ for all four subjects. In Fig. 5 (middle) we plotted “negative surprise” NS versus $P300amp$ for context $C = 1$ and the same for context $C = 2$ in Fig. 5 (right).

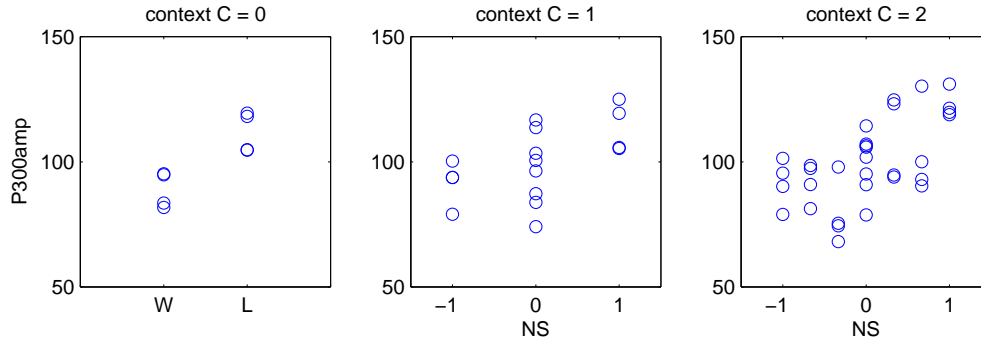


Fig. 5. Left: $P300amp$ for wins (W) and losses (L); Middle: $P300amp$ versus NS for $C = 1$; Right: $P300amp$ versus NS for $C = 2$.

As can be seen from Fig. 5 (left) $P300amp$ for losses was larger than $P300amp$ for wins for all four subjects. The Pearson correlation between $P300amp$ and NS for context $C = 1$ was $corr_1 = .56$ and for context $C = 2$ it was $corr_2 = .60$. Both correlations were significantly different from zero at the five percent error level:

$$t = \frac{.56}{\sqrt{\frac{1-.56^2}{16-2}}} = 2.5291 > t_{95}(14) = 2.1448$$

$$t = \frac{.60}{\sqrt{\frac{1-.60^2}{32-2}}} = 4.1079 > t_{95}(30) = 2.0423$$

4 Discussion

The P300 waveform is one of the most widely studied endogenous evoked potentials (see e.g. [Altenmueller & Gerloff 1999] for a standard text book treatment). The most common paradigm for eliciting a P300 is to randomly intersperse infrequent and therefore unexpected stimuli among frequent stimuli presented to an attentive subject. This is known as the “oddball paradigm”. The amplitude of the P300 is inversely related to the stimulus probability. P300 amplitude is also modulated by subjective outcome probability in an inverse relationship. The latency of the P300 seems to be related to task difficulty. The topography of the P300 reveals a maximum over midline frontal to centro-parietal regions. These aspects have been integrated into a theory of context updating by [Donchin & Coles 1988]. Despite the large amount of work done on the P300 there are still considerable controversies about its functional significance and generators (see e.g. [Verleger 1988] for a critique of the context updating hypothesis and the discussion continuing for more than a decade [Donchin & Coles 1998]). Recently, [Makeig et al. in press] identified at least nine classes of sources of ongoing EEG that contributed to the average P300 response in a visual attention experiment.

The subjective expectancy of a stimulus has been modeled as the sum of a decaying contribution of previously presented like stimuli (memory M), the expectancy for the continuation of alternating pattern (A) and the prior probabilities of the stimuli (P) [Squires et al. 1976]. The amplitude of P300 has repeatedly been found to be inversely related to this estimate of expectancy. Expectancy for the continuation of alternating patterns A has been found to have only a small influence on this relation (only 5 % of accounted variance). The prior stimuli probabilities P are equal for our experiment since overall subjects win and lose equally often during a full game session as expected from game theory ⁸. The memory term M is equal to our term $\hat{E}_{(t)}$ if we set $E_{(i)} = 1$ if $E_{(i)} = E_{(t)}$ and $E_{(i)} = 0$ if $E_{(i)} \neq E_{(t)}$. Therefore memory M is maximal and equals 1 if an expectation is near certain and is fulfilled (e.g. for context $C = 2$ if win-win-win or loss-loss-loss). It is minimal and equals 0 if an expectation is near certain and it is not fulfilled (e.g. for context $C = 2$ if win-win-loss or loss-loss-win). Contrary to our measure of “negative surprise” NS , memory M does not take into account the difference between the stimuli. If we neglect terms A and P , for reasons mentioned above, and correlate memory M with our estimate $P300amp$ we get a correlation of $-.21$ for context $C = 1$ and $.17$ for $C = 2$. Given that we would have expected large negative correlations according to the classical model, it seems that our model of “negative surprise” is better suited to explaining the data. After all, the feedback stimuli had a meaning for our subjects in the context of the “Same/Different”

⁸ This is called Nash-equilibrium, see [Nash 1950].

game and it therefore seemed to make a difference whether our subjects were surprised for the worse or for the better.

Most neuroscientific research on game playing is being carried out within the new emerging field of neuroeconomics. Neuroeconomics tries to correlate mathematical frameworks describing behavioral experiments with physiological measurements from the brain. The mathematical models employed usually are taken from game theory [Von Neumann & Morgenstern 1944] and the goal is to find the neural basis of game theory or, more generally, decision making. A number of projects on game playing monkeys using implanted electrodes have successfully linked single cell firing rates to expected utility. A study on the function of the lateral brain by [Platt & Glimcher 2003] serves as a good example for this line of research. It has long been debated as whether the lateral brain is concerned with sensory, motor or some kind of integrative function. In a series of experiments monkeys had to attend to visual stimuli by looking towards target cues to receive a juice reward. Through analysis of neural spike signals from implanted electrodes, it was shown that neurons in area LIP encode the probability that a movement may yield a reward. This is known and modeled in game theory as relative expected utility. These results cannot directly be linked to our work reported here since we did not yet try to find a connection between the context of the game and movements (button presses in our experiment). However, expected utility is usually expressed as the sum of products of expectations and their values for the player. Expected utility is therefore a weighted sum of expectation terms like our $\hat{E}_{(t)}$. Therefore an approach relating parameters from EEG to the expected utility of movements and game strategic decisions seems to be possible in principle.

There are only few neuroscientific studies of social interaction engaging human subjects in game playing. Most closely related to our approach is probably an fMRI study using the children's game "Paper/Rock/Scissors" [Paulus et al. 2004] which is formally akin to a game of matching pennies with three sided coins. The authors use a temporal difference model to explain their subjects' behavior when playing against a computer. Over the course of the experiment they changed the prior probabilities of the computer playing "Rock", "Paper" or "Scissors". The main result was that a trend process derived from the temporal difference model accurately tracked these changes in prior probabilities and was time-locked to the hemodynamic changes in the inferior frontal gyrus. Contrary to this study, the prior probabilities for our "Same/Different" game were not systematically changed. They rather evolved around the theoretical equilibria due to the subjects' behavior.

[McCabe et al. 2001] had their subjects play a "Trust"-game while recording their fMRI. Their study focused on differences between trials in which subjects played against humans versus trials in which they played against a computer. [Rilling et al. 2002] report differences between cooperative and

non-cooperative phases during games of the so-called “Prisoner’s Dilemma”, monitored using fMRI with only one player being scanned. There are even fewer studies that try to simultaneously monitor brain function of two subjects engaged in any form of social interaction. [Montague et al. 2002] performed simultaneous fMRI in different scanners with pairs of individuals competing against each other in a simple game. The game was designed to measure the effect of deception in a competitive context. Their study is rather a proof of principle experiment than a thorough, neuroscientific study of game playing in humans.

To our knowledge, no one has yet looked into human game playing using simultaneous EEG measurements. In our study of two-person game playing, we have successfully linked the context of the gaming situation to parameters derived from average evoked potentials. This can be seen as a first step toward extending results obtained for single cell firing rates of primates and fMRI measurements of humans to EEG in humans. Average evoked potentials in different game contexts account only for a small part of the event-related brain dynamics one could study. [Makeig et al. 2004] give a systematic view of how to use ICA to go beyond simple response averaging). Given that the midline fronto-central independent components we described are the largest non-artifact components, they nevertheless seemed to be an important part of the overall story.

References

- [Altenmueller & Gerloff 1999] Altenmueller E.O., Gerloff C.: Psychophysiology and the EEG, in: Niedermayer E., Lopes da Silva, F. (eds.): *Electroencephalography: Basic Principles, Clinical Applications, and Related Fields*, Williams & Wilkins, 4th edition, pp. 635-655, 1999.
- [Bell & Sejnowski 1995] Bell A.J., Sejnowski T.J.: An information maximisation approach to blind separation and blind deconvolution, *Neural Computation*, 7(6), pp.1129-1159, 1995.
- [Camerer 2004] Camerer C.F.: *Behavioral Game Theory*, Princeton University Press, Princeton, New Jersey, 2003.
- [Comon 1994] Comon P.: Independent Component Analysis - a new concept?, *Signal Processing*, 36:287-314, 1994.
- [Delorme & Makeig 2004] Delorme A., Makeig S.: EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics, *J Neurosci Methods*, 134:9-21, 2004.
- [Donchin & Coles 1988] Donchin E., Coles M.G.H.: Is the P300 component a manifestation of context updating? *Behavioral and Brain Sciences*, 11:357-374, 1988.

- [Donchin & Coles 1998] Donchin E., Coles M.G.H.: Context updating and the P300, Behavioral and Brain Sciences, 21:152-154,1998.
- [Glimcher 2003] Glimcher P.: Decisions, Uncertainty and the Brain, MIT Press, Cambridge, Massachusetts, London, England, 2003.
- [Hyvaerinen 1999] Hyvaerinen A.: Survey on Independent Component Analysis, Neural Computing Surveys, Vol. 2, 94-128, 1999.
- [Jung et al. 1998] Jung T.-P., Humphries C., Lee T.-W., Makeig S., McKeown M.J., Iraguid V., Sejnowski T.J.: Extended ICA Removes Artifacts from Electroencephalographic Recordings, in Jordan M.I. et al., Advances in Neural Information Processing Systems 10 (NIPS'98), MIT Press/Bradford Books, Cambridge/London, p.894, 1998.
- [Jung et al. 1999] Jung T.-P., Makeig S., Westerfield M., Townsend J., Courchesne E., Sejnowski T.J.: Analyzing and Visualizing Single-Trial Event-Related Potentials, in Kearns M.S. et al. (eds.), Advances in Neural Information Processing Systems 11, MIT Press, 1999.
- [Jung et al. 2000] Jung T.-P., Makeig S., Humphries C., Lee T.-W., Mckeown M.J., Iragui V., Sejnowski T.J.: Removing electroencephalographic artifacts by blind source separation, Psychophysiology, 37, 163-178, 2000.
- [Jung et al. 2001] Jung T.-P., Makeig S., Westerfield M., Townsend J., Courchesne E., Sejnowski T.J.: Analysis and Visualization of Single-Trial Event-Related Potentials, Human Brain Mapping, 14:166-185, 2001.
- [Jutten & Herault 1991] Jutten C., Herault J.: Blind separation of sources I. An adaptive algorithm based on neuromimetic architecture, Signal Processing, 24:1-10, 1991.
- [Lee 1998] Lee T.-W.: Independent Component Analysis: Theory and Applications, Kluwer Academic Publishers, 1998.
- [Makeig et al. 1996] Makeig S., Bell A.J., Jung T-P., Sejnowsky T.J.: Independent Component Analysis of Electroencephalographic Data, in Touretzky D.S. et al. (eds.), Advances in Neural Information Processing Systems 8 (NIPS'96), MIT Press, Cambridge/Boston/London, 145-151, 1996.
- [Makeig et al. 1997] Makeig S., Jung T-P., Bell A.J., Ghahremani D., Sejnowski T.J.: Blind Separation of Auditory Event-related Brain Responses into Independent Components, Proc. of National Academy of Science USA, 94:10979-10984, 1997.
- [Makeig et al. 1999a] Makeig S., Westerfield M., Townsend J., Jung T-P., Courchesne E., Sejnowski T. J.: Functionally independent components of the early event-related potential in a visual spatial attention task, Philosophical Transactions of the Royal Society: Biological Sciences 354:1135-44, 1999.
- [Makeig et al. 1999b] Makeig S., Westerfield M., Jung T-P., Covington J., Townsend J., Sejnowski T. J., Courchesne E.: Functionally independent components of the

- late positive event-related potential during visual spatial attention, *J Neurosci* 19:2665-2680 (1999)
- [Makeig et al. 2002] Makeig S., Westerfield M., Jung T-P., Enghoff S., Townsend J., Courchesne E., Sejnowski T.J.: Dynamic brain sources of visual evoked responses, *Science*, 295:690-694, 2002.
- [Makeig et al. 2004] Makeig S., Debener S., Onton J., Delorme A.: Mining event-related brain dynamics, *Trends in Cognitive Sciences*, Volume 8, Issue 5, pp. 204-210, 2004.
- [Makeig et al. in press] Makeig S., Delorme A., Westerfield M., Townsend J., Courchesne E., Sejnowski T.: Electroencephalographic brain dynamics following visual targets requiring manual responses, *PLOS Biology*, in press.
- [McCabe et al. 2001] McCabe K., Houser D., Ryan L., Smith V., Trouard T.: A functional imaging study of cooperation in two-person reciprocal exchange, *PNAS*, Vol.98, No.20, pp.11832-11835, 2001.
- [Montague et al. 2002] Montague, P.R., Berns, G.S., Cohen, J.D., et al.: Hyperscanning: simultaneous fMRI during linked social interactions, *NeuroImage*, 16(4): 1159-1164, 2002.
- [Nash 1950] Nash J.F.: Equilibrium points in n-person games, *Proc. Nat. Acad. Sciences*, 36, 48-49.
- [Platt & Glimcher 2003] Platt M.L., Glimcher P.W.: Neural correlates of decision variables in parietal cortex, *Nature*, 400:233-38, 1999.
- [Paulus et al. 2004] Paulus M.P., Feinstein J.S., Tapert F.S., Liu T.T.: Trend detection via temporal difference model predicts inferior prefrontal cortex activation during acquisition of advantageous action selection, *NeuroImage* Volume 21, Issue 2 , pp. 733-743, 2004.
- [Rilling et al. 2002] Rilling J., Gutman D., Zeh T., Pagnoni G., Berns G., Kilts C.: A neural basis for social cooperation, *Neuron*, 35(2):395-405, 2002.
- [Roberts & Everson 2001] Roberts S., Everson R.: Independent Component Analysis: principles and practice, Cambridge University Press, Cambridge, UK, 2001.
- [Squires et al. 1976] Squires K.C., Wickens C., Squires N.K., Donchin E.: The Effect of Stimulus Sequence on the Waveform of the Cortical Event-Related Potential, *Science*, Vol. 193, No. 4258, 1142-1146, 1976.
- [Sutton et al. 1965] Sutton S., Braren M., Zubin J., John E.R.: Evoked potentials correlates of stimulus uncertainty, *Science*, 150, 1187-1188, 1965.
- [Verleger 1988] Verleger R.: Event-related potentials and cognition: A critique of the context updating hypothesis and an alternative interpretation of P3, *Behavioral and Brain Sciences*, 11:343-427, 1988.
- [Von Neumann & Morgenstern 1944] Von Neumann J.V., Morgenstern O.: *Theory of Games and Economic Behavior*, Princeton, NJ, Princeton University Press, 1944.