Biased odds for heads or tails: Outcome-evoked P3 depends on frequencies of guesses

ROLF VERLEGER,[a,b] DARIUSZ ASANOWICZ,[a,c] LUCAS WERNER,[a] AND KAMILA ŚMIGASIEWICZ[a]

[a]Department of Neurology, University of Lübeck, Lübeck, Germany
[b]Institute of Psychology II, University of Lübeck, Lübeck, Germany
[c]Institute of Psychology, Jagiellonian University in Kraków, Kraków, Poland

Abstract

Events that had to be predicted evoke large P3 components of the event-related EEG potential. There is conflicting evidence whether these P3s are moderated by participants' preceding guesses. In the present study, participants made one prediction frequently and the other rarely because one stimulus was presented frequently and the other rarely. Thereby, effects on stimulus-evoked P3s of both guess frequency and stimulus frequency could be tested. Indeed, P3s were not only larger with rare than frequent stimuli but also larger after rare than frequent guesses. This result pattern may have additionally been affected by expectancies for payoff. In any case, the modification of outcome-evoked P3 by what had been guessed may reflect that each of the four guess-stimulus combinations is encoded as a separate event category. In terms of the stimulus-response link hypothesis of P3b, it is suggested that P3s are evoked by these events because internal responses (right or wrong) are associated to each of these event categories and need to be reactivated with rare guess-stimulus combinations.

Descriptors: P300, Feedback, Guessing, P3b

The P3 component of ERPs was first reported in a prediction task (Sutton, Braren, Zubin, & John, 1965). Two stimuli (flashes and clicks) were presented in random order. When participants had to predict which stimulus would appear, occurrence of the stimuli evoked a large slow positive potential, soon afterwards termed P3 (Ritter, Vaughan, & Costa, 1968, referring to Davis, Mast, Yoshie, & Zerlin, 1966) or P300 (Donchin & Cohen, 1969). From the very outset, a question of interest was whether this outcome-evoked P3 contains information about participants' preceding guess. Sutton et al. (1965) had already studied whether P3 differs between incorrectly and correctly predicted stimuli and obtained larger amplitudes after incorrect predictions. However, the opposite result, larger P3 amplitudes after correct than after incorrect predictions, was obtained by Levit, Sutton, and Zubin (1973), and no difference was obtained in other studies (Kotchoubey, Grözinger, Kornhuber, & Kornhuber, 1997; Verleger & Cohen, 1978).

A similar diversity of results has been obtained in gambling tasks where participants had to select one of several visible stimuli (symbolized, e.g., in the Iowa Gambling Task as covert decks of cards; Bechara, Damasio, Damasio, & Anderson, 1994) in order to gain or lose some amount of money. The P3 evoked by revealing the value assigned to the chosen stimulus was sometimes larger for losses than wins (Hajcak, Holroyd, Moser, & Simons, 2005; Hajcak, Moser, Holroyd, & Simons, 2007), frequently larger for wins than losses (Bellebaum, Polezzi, & Daum, 2010; Bismark, Hajcak, Whitworth, & Allen, 2013; Gentsch, Grandjean, & Scherer, 2013; Kimura & Katayama, 2013, as a tendency; Wu & Zhou, 2009; Yang, Guo, Tang, & Luo, 2013), and sometimes did not differ (Goyer, Woldorff, & Huettel, 2008; Yeung & Sanfey, 2004). Thus, as already conjectured by Sutton et al. (1965, p.150), it appears that success or failure of guessing per se does not affect P3. Rather, this effect might interact with other factors, including stimulus probabilities, payoff structure, physical differences between stimuli, and sequence of preceding stimuli (cf. Munson, Ruchkin, Ritter, Sutton, & Squires, 1984).

When two stimuli are presented in unpredictable random series, one stimulus frequently and the other rarely, then the rarely occurring stimuli will evoke larger P3 amplitudes than the frequent ones (oddball effect; Duncan-Johnson & Donchin, 1977; Squires, Squires, & Hillyard, 1975). In the present study, the two stimuli to be guessed differed in their probabilities (80% vs. 20% of trials). Participants were informed about the frequency relation by both oral and written instructions. Therefore, their guesses are expected to reflect these probabilities: Ideally, *frequent* guesses will occur in 80% of trials and *rare* guesses in 20%. This yields expected probabilities of $0.80 \times 0.80 = 64\%$ for the f → F combination (guess the

frequent outcome, outcome is indeed the frequent one), $0.20 \times 0.80 = 16\%$ for r → F (guess rare, but outcome is frequent), $0.80 \times 0.20 = 16\%$ for f → R, and $0.20 \times 0.20 = 4\%$ for r → R. This distribution of events provides an opportunity to test whether outcome-evoked P3s contain traces of participants' preceding guesses, because there are not only frequent and rare stimuli but, independently, frequent and rare guesses. The question is whether rare guesses will lead to increased amplitudes of P3s evoked by the guessed outcomes. This question was already pursued by Tueting, Sutton, and Zubin (1970) who indeed found that P3 increased with decreasing frequencies both of outcomes and of guesses. But data were reported from four participants only and recordings were made from frontocentral sites only, thereby possibly missing the bulk of the P3 complex, which, in active tasks, consists of a small frontocentral P3a and a large centroparietally focused P3b (Debener, Makeig, Delorme, & Engel, 2005; Dien, Spencer, & Donchin, 2004; Gaeta, Friedman, & Hunt, 2003; Verleger, Jaśkowski, & Wascher, 2005).

Therefore, this question was addressed anew in the present study, stating the following hypotheses. Undoubtedly, due to the oddball effect, outcome-evoked P3 will be affected by the 80/20 stimulus probability: rare stimuli will evoke larger P3 amplitudes than frequent stimuli (R > F). This will be the only effect if guess frequency is not reflected in the P3 evoked by guessed events. In contrast, we assume that also rare guesses will lead to larger outcome-evoked P3 amplitudes than frequent guesses (r > f), thereby causing the outcome-evoked P3 to contain information about the preceding prediction. Combining this presumed effect of prediction with the uncontroversial effect of stimulus probability: P3 is expected to be largest for r → R, smaller for r → F and f → R, and smallest for f → F, reflecting the inverse frequencies of these guess-outcome combinations (4%, 16%, 16%, 64%). This combination implies opposite effects of guessing correctly for frequent and rare stimuli: P3 amplitudes will be larger for incorrectly than correctly guessed frequent stimuli (r → F > f → F, because 16% is less frequent than 64%) but smaller for incorrectly than correctly guessed rare stimuli (f → R < r → R, because 16% is more frequent than 4%).

There is the risk that participants would not truly guess but rather press any key, to pass through the task as quickly as possible. Moreover, even if participants are cooperative enough to play the guessing game without any external reward, guided by internal motivation, then, if this internal reinforcement is provided by guessing correctly as much as possible, the optimum strategy consists of always guessing the frequent event. Therefore, to achieve the desired 80/20 distribution of guesses, corresponding to the 80/20 distribution of stimuli, we assigned payoffs to rare and frequent stimuli that are inversely proportional to their probabilities. Payoffs of +2 cents (ct.) were assigned for correctly guessed frequent events, of +8 ct. for correctly guessed rare events, and of 0 ct. for incorrect guesses. Thereby, making the frequent guess became as rational as making the rare guess: The expected payoff is equal (80% * 2 ct. = 20% * 8 ct.) and nothing would be lost in case of incorrect predictions. Thus, we hoped to obtain frequent and rare guesses in a ratio of 80/20 corresponding to their objective probabilities. Whether the intended effects of guess probability were moderated by unintended effects of payoff value will be addressed in the General Discussion section.

Effects of the preceding guess on the outcome-evoked P3 might be mediated by processes reflected by electrophysiological events measurable between guessing and outcome. Because outcome-evoked P3 is measured against a baseline before outcome onset,

any increase of outcome-evoked P3 after rare guesses might actually be due to increased levels of negativity of this prestimulus baseline after such rare guesses. To control for this possibility, the stimulus-preceding negativity (SPN) was measured, which developed in the 1-s interval between guessing and outcome onset (cf. Kotchoubey et al., 1994).

In order to distinguish effects of guessing more clearly from the act of simply pressing a key, the mapping of guesses to the two keys was varied in Experiment 1. In *fixed* blocks, key assignments were constant, for example, pressing left always meant guessing the letter X, and pressing right meant guessing the letter U. In *alternating* blocks, assignments were *alternating* between trials. Experiment 2 served as control experiment to corroborate the major result of Experiment 1.

Experiment 1

Experiment 1 aimed at studying effects of making frequent or rare guesses on P3 amplitudes evoked by the guessed stimuli. We expected that making the rare guess would be reflected in larger P3s evoked by the outcomes, independently of outcome probabilities.

To control for the possibility that participants mechanically pressed some key instead of guessing, how guesses were mapped to the two keys was varied. In fixed blocks, key assignments were constant across trials. In alternating blocks, assignments were alternating between trials.

Method

Participants. Sixteen right-handed university students participated (8 females and 8 males, aged 24 years on average, $SD = 2.2$). All participants reported normal or corrected-to-normal vision and no history of neurological disorders. Informed written consent was obtained before the experiment, and participants were paid depending on their success in guessing, amounting to 15€–17€, paid out after the experiment. (See below for details).

Stimuli and procedure. Participants were seated in a comfortable armchair in a darkened room, with about 1.1 m viewing distance from the 17" computer screen, and held a computer keyboard on their lap. They were informed by written instruction on the screen that the task was a gambling game requiring some luck, that they had to guess which of the two letters X or U would occur, and that one letter would occur frequently and the other rarely. Accurate guesses would yield 2 ct. for frequent letters and 8 ct. for rare letters, and no money would be lost after inaccurate guesses. The task consisted of 1,000 trials, divided into four blocks. After every 20 trials and at block ending, summary feedback was given, separately for frequent and rare outcomes, on the number of correct guesses and the amount of money earned. Presentation software 14.0 (www.neurobs.com) was used to present the stimuli, register responses, and send stimulus and response codes to another computer, where codes were stored with the recorded EEG.

As illustrated in Figure 1, trials started with a guess prompt below screen center ("guess, please" in German, in 20 pt. font) displaying the two letters left and right below the prompt, as a reminder about which key, left or right, was assigned to which letter. These texts were black on a light gray screen. Participants made their guess by pressing the left or right Ctrl keys on the keyboard. To prevent premature mechanistic guessing, error messages ("pressed too early," in German) appeared in large red 30-pt. font
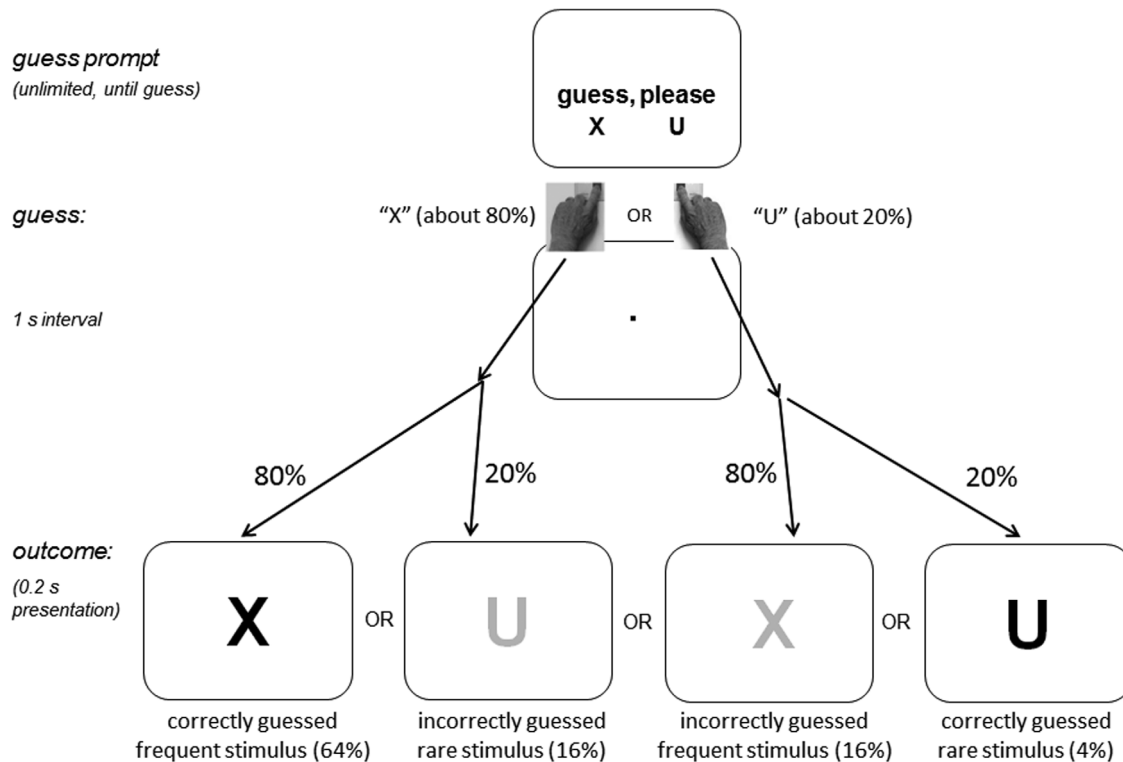
Figure 1. Outline of the paradigm. The guess prompt informed about mapping of the two stimuli to keys (here, X left, U right), which was fixed within a block or (in Experiment 1) alternated between trials. Outcomes appeared 1 s after guessing. In Experiment 1 and in the explicit condition of Experiment 2, blue or yellow denoted right or wrong (here, black and gray). One stimulus was frequent (X), the other rare (U), schematically 80%/20% (in Experiment 1 actually 81%/19%). This proportion was expected to be reflected in participants' guess probabilities, 80%/20%. If so, the four outcomes had probabilities of 64%, 16%, 16%, and 4%.

for 4 s whenever the keys were pressed before onset of the guess prompt. Key pressing blanked the screen and was followed after 1 s by the outcome, which was an X or U (Helvetica, 35 pt.) displayed at screen center for 200 ms. X was the frequent letter for half the participants, and U for the other half. The outcome letter was displayed in color, providing explicit feedback about guess accuracy. For half the participants, blue meant positive feedback and yellow meant negative feedback, and vice versa for the other half (crossed with assigned frequencies of X vs. U). The guess prompt of the next trial was presented 1 s later.

The mode of assigning keys to stimuli was varied between blocks. Within the two "fixed" blocks, assignment remained constant: In one block, X was guessed by pressing left and U by pressing right, and vice versa in the other block. In the two "alternating" blocks, assignment alternated with each trial. Four different block orders were used across participants (crossed with the above balancing of letter frequency and color meaning): Fixed blocks were either blocks #1 and #3, or #2 and #4, and of these two fixed blocks, either the left-frequent or the right-frequent block was presented first.

The letter to be presented was randomly chosen in each trial with 80/20 probabilities. An additional restriction against immediate repetitions of two rare letters was introduced. In this case, the frequent letter was presented instead of the repeated rare letter, which was deferred to the following trial (where its probability was, thereby, $0.2 + 0.2 - 0.2*0.2 = 36\%$). This restriction diminished the global probability of rare letters from 20% to 19.2%.

Analysis of behavior and of outcome percentages. Analyses of variance (ANOVAs) on repeated measurements were used for

statistical analyses. The major behavioral measure was given by the frequencies of the four guess-outcome combinations, which resulted from the objective probabilities of events and from participants' guessing behavior. To analyze outcome frequencies, percentages of correctly and incorrectly guessed frequent and rare events were determined in either experimental condition (fixed vs. alternating key mapping) and were analyzed in an ANOVA with the factors outcome ($f \rightarrow F$, $r \rightarrow F$, $f \rightarrow R$, $r \rightarrow R$) and key assignment (fixed vs. alternating) and in ensuing pairwise comparisons between outcomes. Percentages of guessing the rare events were compared to actual percentages in an ANOVA with the factors subjective/objective (guessing vs. actual frequency) and key assignment (as before).

Additionally, the speed of guessing was measured, by response times (RTs) of pressing the keys (for guessing) in response to the guess prompt. Averaged RTs were entered to an ANOVA with the factors frequency of guessed event (frequent vs. rare) and key assignment (fixed vs. alternating).

EEG recording and analysis. EEG was recorded with Ag/AgCl electrodes (Easycap, www.easycap.de) from 60 scalp sites, including 8 midline positions from AFz to Oz and 26 pairs of symmetric left and right sites. Additional electrodes were placed at the nose tip for offline reference and at Fpz as connection to ground. Online reference was Fz. For artifact control, electrooculogram (EOG) was recorded, vertically (vEOG) from above versus below the right eye and horizontally (hEOG) from positions next to the left and right tails of the eyes. Data were amplified from DC to 250 Hz by a BrainAmp MR plus and stored at 500 Hz per channel. Offline
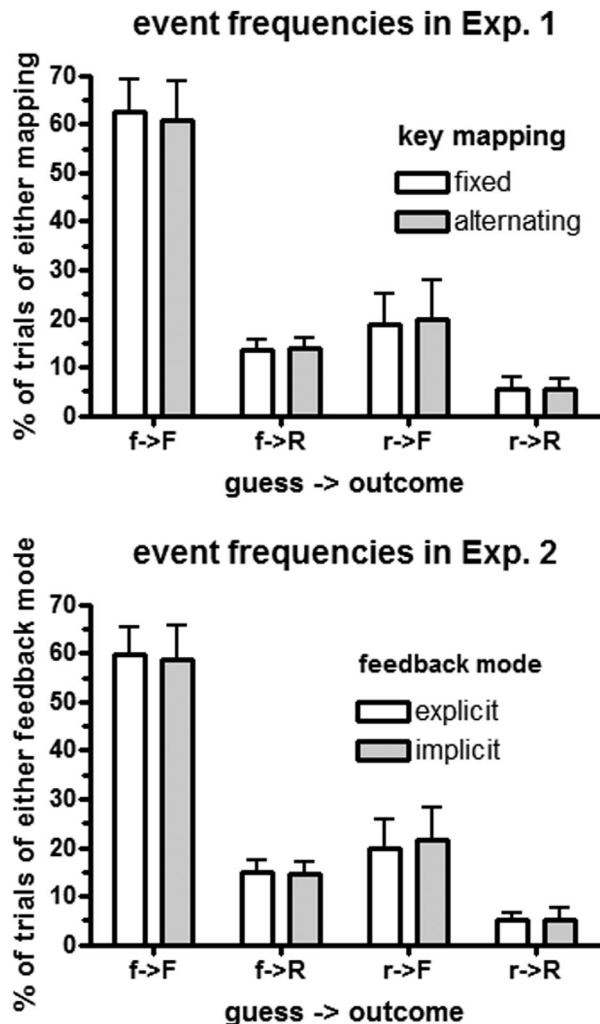
event frequencies in Exp. 1

event frequencies in Exp. 2

Figure 2. Mean frequencies (in %) for each guess-outcome combination. f and r on *x* axis denote frequent and rare guesses, F and R denote frequent and rare outcomes. Open and filled bars represent the fixed and alternate key assignment conditions of Experiment 1 (upper), and the explicit and implicit feedback on guess correctness of Experiment 2 (lower). (Open-bar conditions remained equal across the two experiments). The order of conditions from left to right is the same as in Figure 1.

processing was done with BrainVision Analyzer software (version 2.03). Data were rereferenced to the nose tip, low-pass filtered at 25 Hz, and segmented to epochs of 2.1 s duration, from 100 ms before the guess key press to 1 s after outcome onset. These epochs were edited for artifacts. First, epochs were rejected as gross artifacts when consecutive data points differed by more than 50 μV (except EOG and AF3, AFz, AF4, lest trials would be rejected for blinks). Then, ocular artifacts were corrected by using the linear regression method implemented in the BrainAnalyzer software. Finally, data were referred to the mean amplitude of the first 100 ms before outcome onset (i.e., the 900–1,000 ms interval after the guess response) as baseline in each channel, and trials were rejected when voltages exceeded ± 150 μV in any EEG channel. Mean numbers of remaining trials included for analysis were, for fixed and alternating mapping, respectively, 291 and 286 for the f → F combination (minima: 242 and 191), 90 and 96 for r → F (minima: 40 and 49), 62 and 64 for f → R (minima: 45 and 45), and 26 and 26 for r → R (minima: 9 and 13).

EEG data were then averaged over trials, separately for the four guess-outcome combinations in the two key-assignment conditions. Parameters were measured in these averaged waveforms. After inspecting these waveforms and their topographic distributions, the latency window for searching the P3 peak was defined as 250–650 ms after letter onset at the seven midline sites, Fz, FCz, Cz, CPz, Pz, POz, Oz. P3 amplitudes were found to be largest at FCz, Cz, CPz, Pz in a first analysis on all seven midline sites. The main analysis, therefore, used data from those four sites, with the factors recording site (4 levels), stimulus frequency (frequent, rare), guess frequency (frequent, rare), and key assignment (fixed, alternating). Amplitudes of SPN developing in the 1 s between guess and outcome onset were measured as the level of ERPs at 100–0 ms before outcome onset (serving as the baseline for P3 measurement), referred to the level of activity 1 s before (at 100–0 ms before the guess key press). These measurements were taken in averages pooled across frequent and rare events because frequencies of these events could not affect the preceding SPN. SPN amplitudes at the seven midline sites were submitted to ANOVA with the factors recording site (7 levels), guess frequency, and key assignment (as above). A second ANOVA was conducted on the PO3, POz, and PO4 sites where SPN was largest, using the same factors (with 3 levels for the recording site factor).

Degrees of freedom of the recording site factor were corrected with the Greenhouse-Geisser method. To clarify interactions, ANOVAs were conducted on the single levels of the interacting factors. Pearson correlations between frequencies and P3 amplitudes were calculated across participants on the differences f → F minus r → F, r → F minus f → R, and f → R minus r → R, in order to measure whether individual variations in P3 increase were related to individual variations in frequencies of guess-outcome combinations.

Results

Behavior. The top panel of Figure 2 presents the frequencies of the four outcomes under the two key assignments (fixed vs. alternating). These frequencies amounted to 62%, 14%, 19%, and 5% for f → F, f → R, r → F, r → R, respectively, pooled across key assignments. These frequencies differed from each other globally, $F(3,45) = 275.6$, $p < .001$, and in ensuing pairwise analyses, $p = .02$ for r → F vs. f → R, and $p \leq .001$ for all other comparisons. Thus, frequent stimuli were more often guessed correctly than incorrectly (62% vs. 19%) whereas rare stimuli were more often guessed incorrectly than correctly (14% vs. 5%). Frequencies did not differ between fixed and alternating key assignments, Outcome × Key Assignment, $F(3,45) = 2.1$, $p = .14$.

Proportions of frequent and rare events as randomly generated by the control program amounted to 80.9% (f → F and r → F) and 19.1% (r → R and f → R; averages across participants, ± 1.1%) which comes close to the expected probabilities of 80.8% and 19.2% (cf. Method). Proportions of participants' guesses of frequent and rare outcomes amounted to 75.3% (f → F and f → R) and 24.7% (r → F and r → R; ± 8.9%). That is, rare events were predicted more frequently (and, correspondingly, frequent events less frequently) than they occurred, $F(1,15) = 6.3$, $p = .02$. Accordingly, frequent events were successfully predicted in 76.2% (± 8.7%) of their occurrences, f → F/(f → F + r → F), somewhat less than their probabilities, and rare events were successfully predicted in 28.1% (± 10.8%) of their occurrences, r → R/(r → R + f → R), somewhat more than their probabilities.

Participants differed in overall frequencies of rare guesses (varying from 12% to 42%) but, rather than being random, this
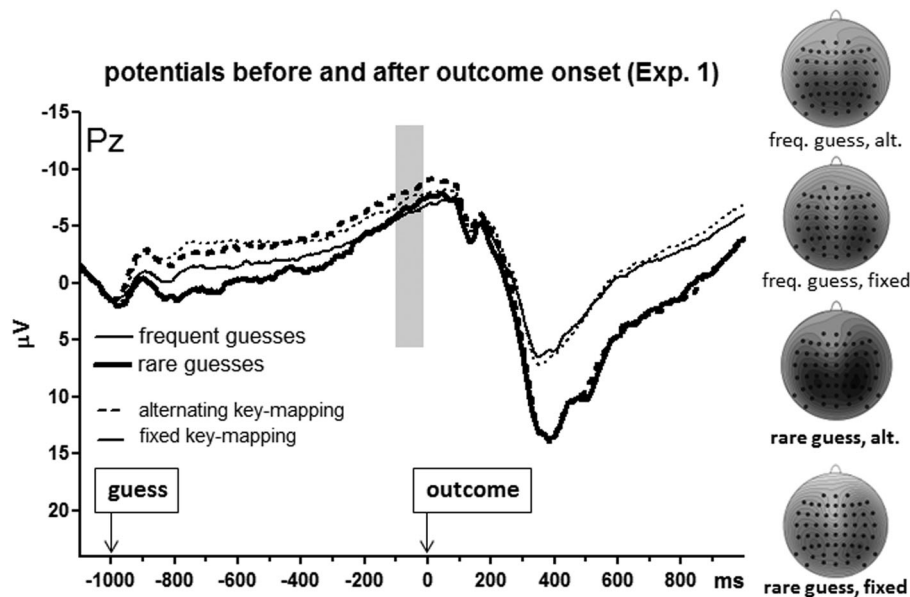
Figure 3. Pz waveforms and scalp topographies of SPN for frequent and rare guesses in Experiment 1. Data have been pooled across frequent and rare outcomes. Waveforms are displayed for frequent and rare guesses (thin vs. bold lines) with fixed-key versus alternating-key assignment (solid vs. dashed lines). Unit of *x* axis is ms, with the zero point denoting outcome onset. Unit of *y* axis is μV, with negative polarity plotted upwards. The maps display scalp topographies at 100 ms before outcome onset (epoch depicted by gray rectangle). View on the head is from above. Each map is scaled from negative maximum (black) to zero (white).

variability reflected quite stable individual response strategies, correlating across participants between fixed and alternating key assignments ($r = .93, p < .001$).

Participants needed more time for making their guess in response to the guess prompt in the alternating-key than in the fixed-key condition (717 ms vs. 565 ms; $F(1,15) = 25.1, p < .001$) and guessing rare tended to take longer than guessing frequent when keys alternated (Key Assignment × Frequency of Guessed Event: $F(1,15) = 6.5, p = .02$; effect of frequency for alternating keys: $F(1,15) = 3.1, p = .10$) but not when keys were fixed (effect of frequency for fixed keys: $F(1,15) = 0.0, n.s.$).

ERPs. *SPN before onset of outcome events.* Figure 3 depicts the time course of ERPs from the moment of the guess until 1 s after outcome onset. Evidently, there is a distinct negative-going stimulus-preceding potential between guessing and outcome onset. This SPN was largest at the back of the head (Pz and POz in midline analysis, main effect of recording site $F(6,90) = 12.3, p < .001$) and larger at lateral sites, in particular PO3 and PO4, than at midline (recording site in the ANOVA on PO3, POz, PO4: $F(2,30) = 4.7, p = .03$). In the present context, the major result is that SPN did not differ significantly after frequent and rare guesses (guess frequency in midline analysis: $F(1,15) = 0.3, n.s.$; Guess Frequency × Recording Site: $F(1,15) = 1.2, n.s$) in stark contrast to outcome-evoked P3 (see below). There was a triple interaction of Guess Frequency × Recording Site × Key Assignment, though, $F(6,90) = 4.0, p = .02$, reflecting effects of Key Assignment × Guess Frequency at Fz and FCz (at either site $F(1,15) = 6.5, p = .02$) because least negative values were reached at these sites when rare guesses were made under fixed key assignment. The only other effect was an interaction of Recording Site × Key Assignment, $F(6,90) = 2.7, p = .046$, because SPNs had their midline maximum slightly less posteriorly (Pz and POz) with alternating assignment than with fixed assignment (POz and Oz). But

amplitudes did not differ between key assignments at any of these midline sites ($p = .08$ at best).

P3 amplitude. Already Figure 3 shows that P3s evoked by the guessed stimuli were larger after rare than frequent guesses. Figure 4 focuses on these stimulus-evoked ERPs, breaking frequent and rare predictions (thin and bold lines, respectively, both in Figure 3 and Figure 4) further down to frequent and rare outcomes (gray and black lines, respectively).

P3 amplitudes were analyzed from the midline sites FCz, Cz, CPz, Pz where P3 was largest (cf. maps in Figure 4). Mean values are compiled in Figure 5. P3 amplitudes were affected not only by stimulus frequency, $F(1,15) = 98.3, p < .001$, with larger amplitudes for rare than frequent stimuli (black vs. gray in Figure 4) but, unlike the preceding SPN, as well by guess frequency, $F(1,15) = 30.1, p < .001$, with larger amplitudes after rare than frequent guesses (bold vs. thin in Figure 4). The interaction of these two factors, $F(1,15) = 8.8, p = .01$, indicated that effects of preceding guesses were larger for frequent than for rare stimuli (effect of guess frequency in separate analysis of frequent stimuli: $F(1,15) = 45.2, p < .001$; for rare stimuli: $F(1,15) = 8.8, p = .009$). Overall, amplitudes were larger at CPz and Cz than at Pz and FCz, $F(3,45) = 5.7, p = .02$. This effect of site interacted with stimulus frequency, $F(3,45) = 6.9, p = .01$, because amplitude maxima shifted from a more posterior focus with the smaller P3 evoked by frequent stimuli (CPz and Pz) towards a more anterior focus with the larger P3 evoked by rare stimuli (Cz and CPz).

Key assignment interacted with Stimulus Frequency × Guess Frequency, $F(1,15) = 7.0, p = .02$, because P3 tended to be larger with alternating than fixed assignment with frequent stimuli (effect of key assignment in separate analysis of frequent stimuli: $F(1,15) = 3.8, p = .07$) and had the same tendency with rare stimuli after rare guesses but the opposite tendency with rare stimuli after frequent guesses (effect of Key Assignment × Guess Frequency for rare stimuli: $F(1,15) = 6.4, p = .02$). Furthermore, key
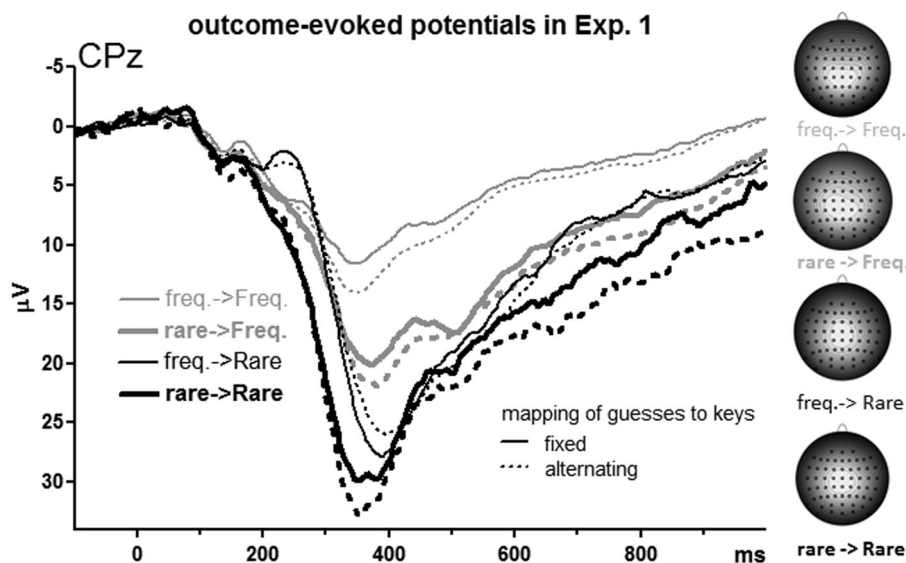
Figure 4. CPz waveforms and scalp topographies of P3 for each guess-outcome combination in Experiment 1. Waveforms are displayed in each of the four combinations: f → F (thin gray), r → F (bold gray), f → R (thin black), and r → R (bold black), separately for fixed (solid) and alternating (dotted) mapping of guesses to keys. Unit of *x* axis is ms, with the zero point denoting outcome onset. Unit of *y* axis is μV, with negative polarity plotted upwards. The maps display scalp topographies at the time points of the P3 peaks (340 ms, 390 ms, 400 ms, 350 ms for f → F, r → F, f → R, r → R, respectively, pooled across the two key assignment conditions). View on the head is from above. Each map is scaled from minimum (black) to maximum (white).

assignment tended to interact with Stimulus Frequency × Site, $F(3,45) = 3.0$, $p = .06$, because with rare stimuli P3 was equally large at Cz and CPz with alternating mapping and somewhat larger at Cz than CPz with fixed mapping.

Correlations between outcome frequencies and P3 amplitudes. To confirm the inverse relationship between P3 amplitudes and event frequencies also in terms of between-subject variations, correlations were computed across participants between decreases in event frequencies and the corresponding increases in P3 amplitudes at CPz, pooled across fixed and alternating key assignments. Indeed, decreases of frequencies correlated with P3 increases (one-sided *p* values): from f → F to r → F $r(14) = -.56$, $p = .01$; from r → F to f → R $r(14) = -.50$, $p = .02$; from f → R to r → R $r(14) = -.42$, $p = .06$.

Discussion

P3 amplitudes evoked by guessed stimuli were not only affected by stimulus frequencies but as well by frequencies of the preceding guesses. This is in line with the hypothesis that P3 evoked by outcome stimuli contains information about participants' preceding guesses.

Key assignment was found to be of minor importance: Results were highly similar with fixed and alternating key assignments, and participants' proportions of guessing the rare stimuli correlated almost perfectly between these two assignment modes. This is good evidence that participants did not mechanically press the keys instead of guessing.

Also of minor importance were RTs of pressing the keys for making the guess in response to the guess prompt: Effects on P3 amplitudes were very similar when keys were fixed and when keys alternated, whereas RTs were generally larger when keys alternated. Further, increases of P3 amplitudes after rare compared to frequent guesses were similar for fixed and alternating conditions,

while RTs of rare and frequent guesses tended to differ from each other only when keys alternated but not when assignment was fixed.

Finally, the effect of frequencies of guesses on outcome-evoked P3 amplitudes was not moderated by parallel effects on the SPN developing between guesses and stimulus onsets. Rather, effects of guess frequency on the SPN were restricted to rare guesses under fixed key assignment. The topography of this effect, at anterior sites near the hand-motor areas, suggests that this effect is related to the fact that the hand making the rare guess is used less frequently during the fixed key-assignment block (while both hands are used equally often with alternating key assignment). Apart from this possible motor contribution, similar to the contingent negative variation (CNV) preceding an imperative signal that requires a motor response (Walter, Cooper, Aldridge, McCallum, & Winter, 1964), SPN had its topographical focus at posterior-lateral sites PO3 and PO4. Any rigorous evaluation of SPN's topography must take into account that its baseline (100 ms before the guess) is probably not electrically neutral (Urbach & Kutas, 2006) but may rather be affected by potentials evoked by the guess prompt, presented 500–700 ms before participants made their guess (cf. analysis of RTs). This said, such posterior-lateral topography, near the visual cortices, does make some sense for this SPN because no hand-motor action was being prepared (underlying the usual frontocentral topography of CNV) but rather a visual stimulus (the guessed outcome) was being expected (cf. Simons, Öhman, & Lang, 1979).

However, before drawing the strong conclusion that P3 evoked by outcome stimuli contains information about participants' preceding guesses, an obvious objection must be considered. In fact, the very outcome stimuli differed depending on the preceding guess: As indicated by the black and gray shades of letters in Figure 1, correctly and incorrectly guessed stimuli had different colors (blue vs. yellow). We had done so as a further safeguard against participants' mechanically pressing some key and not being
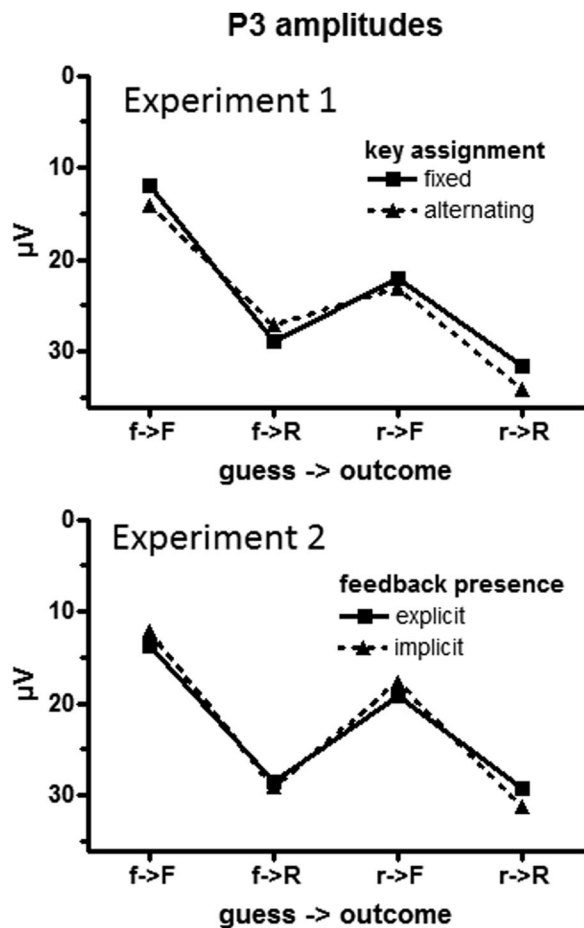
P3 amplitudes

Figure 5. Mean amplitudes (in μV) of the P3 peak for each guess-outcome combination. On *x* axis, f and r denote frequent and rare guesses, F and R denote frequent and rare outcomes. Solid and dashed lines represent fixed and alternate key assignment conditions of Experiment 1 (upper), and explicit and implicit feedback on guess correctness in Experiment 2 (lower). (Solid-line conditions remained equal across the two experiments). *y* axis is inverted, with positive downwards, to resemble polarity conventions of Figures 3, 4, 6, 7.

interested in guess outcomes. Other parts of this motivation package were the money reward for gains and the feedback screens after every 20 trials. But this tainting of letters depending on success implied that, rather than reflecting information about participants' preceding guesses, the increase of P3 after rare guesses in the present task might have simply reflected differences between the evoking outcome stimuli.

Experiment 2

Experiment 2 tested whether the dependency of P3 amplitudes on frequency of the preceding guess was indeed due to participants' integrating the guessed stimuli with their preceding guesses or whether this dependency was only obtained because letter colors differed between correctly and incorrectly guessed stimuli. This explicit cueing of prediction success was now used in one part of Experiment 2 only. Crucially, in the other part, letters were uniformly black such that information about correctness of predictions was only implicitly provided by the outcome stimuli, requiring participants to compare stimuli to preceding predictions in order to know whether predictions were correct.

Second, we wanted to make sure that results of Experiment 1 had not been biased by the restriction on randomness. Therefore, this restriction was removed, thereby allowing rare stimuli to follow each other, which yielded expected probabilities of exactly 80% and 20% for frequent and rare stimuli.

Method

Differences from Experiment 1 will be described only.

Participants. Sixteen right-handed undergraduate students participated (9 females and 7 males, aged 24 years on average, $SD = 3.1$). One other participant had to be replaced because he did not play the guessing game according to our instructions but rather kept on pressing the rare-guess key (perfectly rationally) in about 75% of trials.

Stimuli and procedure. Two of the four blocks were identical to the fixed-key blocks of Experiment 1. This will be called the *explicit* part because letter color provided explicit information about whether the guess was correct. In the *implicit* part, letters were always black. Thereby, participants had to relate the letter to their preceding guess in order to know whether their guess was correct. Either part consisted of two blocks with fixed key assignment, where either the right or the left hand made the frequent guess and the other hand the rare guess. Half the participants had the explicit part first and half the implicit part. Within either part, half the participants had the frequent guess first at the right key and half at the left key.

There was no restriction against immediate repetitions of two rare letters any more. Thus, the probabilities for frequent/rare stimuli were exactly 80/20.

Data analysis. Mean numbers of trials available for EEG analysis after editing for artifacts were, for explicit and implicit feedback, respectively, 282 and 269 for the f → F combination (minima: 205 and 129), 95 and 99 for r → F (minima: 47 and 59), 70 and 66 for f → R (minima: 41 and 37), and 24 and 25 for r → R (minima: 14 and 15).

In the ANOVAs, the key assignment factor was replaced by the feedback type factor (explicit, implicit).

Results

Behavior. Frequencies of the four outcomes under the two feedback conditions are presented in the bottom panel of Figure 2. There was no difference between explicit and implicit feedback, Outcome × Feedback Type $F(3,45) = 1.1$, *n.s.* Pooled across feedback types, frequencies amounted to 59%, 15%, 21%, and 5% for f → F, f → R, r → F, r → R, respectively. These frequencies differed from each other globally, $F(3,45) = 374.5$, $p < .001$, and in ensuing pairwise analyses, $p = .006$ for r → F versus f → R, and $p < .001$ for all other comparisons.

As intended, proportions of frequent and rare stimuli, randomly generated by the control program, amounted to 80.0% (f → F and r → F) and 20.0% (r → R and f → R; averages across participants, ±1.1%). Proportions of participants' guesses of frequent and rare stimuli amounted to 74.0% (f → F and f → R) and 26.0% (r → F and r → R; ± 7.2%). Thus, like in Experiment 1, rare stimuli were predicted more frequently (and, correspondingly, frequent stimuli less frequently) than they occurred, $F(1,15) = 10.3$, $p = .006$. Accordingly, frequent stimuli were successfully predicted in 74.1%
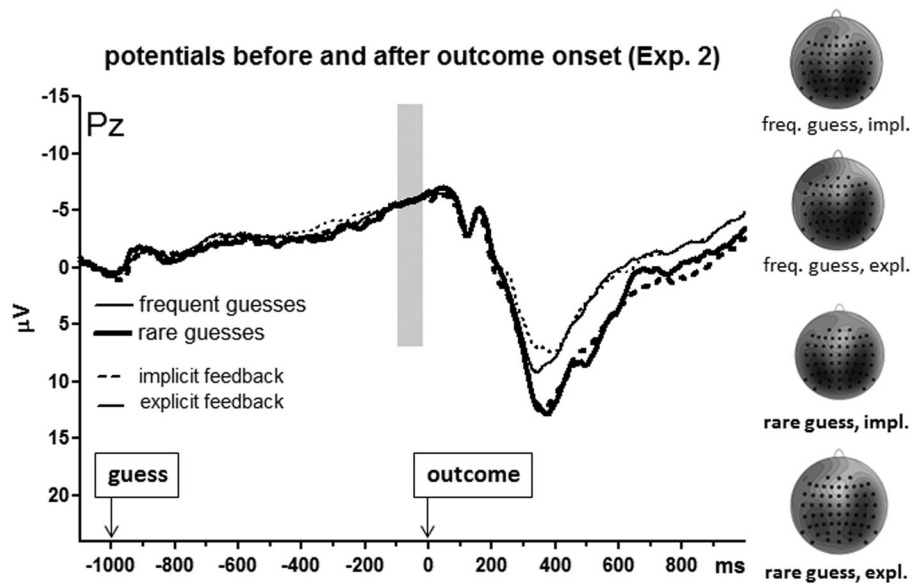
Figure 6. Pz waveforms and scalp topographies of SPN for frequent and rare guesses in Experiment 2. Data have been pooled across frequent and rare outcomes. Waveforms are displayed for frequent and rare guesses (thin vs. bold lines) with explicit versus implicit feedback (solid vs. dashed lines). Unit of *x* axis is ms, with the zero point denoting outcome onset. Unit of *y* axis is μV, with negative polarity plotted upwards. The maps display scalp topographies at 100 ms before outcome onset (epoch depicted by gray rectangle). View on the head is from above. Each map is scaled from negative maximum (black) to zero (white).

($\pm$ 7.2%) of their occurrences, f → F/(f → F + r → F), somewhat less than their probabilities, and rare stimuli were successfully predicted in 26.2% ($\pm$ 7.8%) of their occurrences, r → R/(r → R + f → R), somewhat more than their probabilities.

Participants differed in their overall frequencies of rare guesses (varying from 16% to 45%), which, again, reflected quite stable individual response strategies, correlating across participants between explicit and implicit feedback blocks ($r = .73$, $p = .001$).

Times for making the guess in response to the guess prompt varied around 600 ms and remained unaffected from feedback type and guess frequency, $F(1,15) \leq 0.9$, *n.s.*

ERPs.

SPN before onset of outcome events. Figure 6 depicts the time course of ERPs from the guessing key press until 1 s after outcome onset. Like in Experiment 1, SPN was largest at the back of the head (POz in midline analysis), main effect of recording site $F(6,90) = 13.1$, $p < .001$. Neither guess frequency ($F \leq 1.7$, $p \geq .21$) nor feedback type ($F \leq 1.4$, $p \geq .23$) had significant effects. SPN amplitudes were numerically larger at lateral sites PO3 and PO4 than at POz but this difference was not significant, unlike in Experiment 1 (recording site: $F(2,30) = 2.1$, $p = .14$) nor were there other significant effects in this analysis on PO sites ($F \leq 2.5$, $p \geq .12$).

P3 amplitudes. Grand-average ERP waveforms evoked by the outcomes are shown in Figure 7, and mean amplitudes are compiled in the lower panel of Figure 5.

Very similar to Experiment 1, P3 amplitudes were affected both by stimulus frequency, $F(1,15) = 123.1$, $p < .001$, with larger amplitudes for rare than frequent stimuli (black vs. gray in Figure 7), and by guess frequency, $F(1,15) = 23.7$, $p < .001$, with larger amplitudes after rare than frequent guesses (bold vs. thin in Figure 7). These two factors interacted, $F(1,15) = 8.9$, $p = .009$, indicating that large

effects of guesses occurred for frequent stimuli, $F(1,15) = 43.8$, $p < .001$, for the main effect of guess frequency in separate analysis of frequent stimuli, but not for rare ones, $F(1,15) = 1.8$, *n.s.*, in separate analysis of rare stimuli. Overall, amplitudes were largest at CPz and smallest at FCz, $F(3,45) = 5.5$, $p = .02$, without interaction with any other factor, $F(3,45) \leq 1.7$, $p \geq 0.21$.

Importantly, feedback type did not modify effects of guess frequency, all interactive effects of this combination of factors $F \leq 1.1$, $p \geq .36$. The one effect of the feedback factor was its interaction with stimulus frequency, $F(1,15) = 7.2$, $p = .02$, reflecting that P3 tended to be smaller with implicit than explicit feedback for frequent stimuli, $F(1,15) = 3.2$, $p = .10$, but not for rare stimuli, $F(1,15) = 1.1$, *n.s.*

Correlations between outcome frequencies and P3 amplitudes. Correlations across participants between decreases in event frequency and the corresponding increases in P3 amplitude at CPz, pooled across explicit and implicit feedback, were significant for the difference between f → F and r → F, $r(14) = -.55$, $p = .01$ (one-sided), but not between r → F and f → R, $r(14) = -.22$, *n.s.*, nor between f → R and r → R, $r(14) = -.20$, *n.s.* When computing these correlations for explicit and implicit blocks separately, it was found that frequencies and P3 correlated throughout for explicit feedback: $r(14) = -.49$, $p = .03$, for f → F – r → F; -.34, $p = .10$, for r → F – f → R; -.50, $p = .03$, for f → R – r → R. In contrast, for implicit feedback, only the first large difference between f → F and r → F produced this correlation: $r(14) = -.51$, $p = .02$, for f → F – r → F; -.10, *n.s.*, for r → F – f → R; -.07, *n.s.*, for f → R – r → R.

Discussion

P3 amplitudes were larger for incorrectly than correctly predicted frequent stimuli. This influence of the preceding guess was independent from explicitness of feedback, that is, from whether incorrectly and correctly predicted stimuli differed by their color or
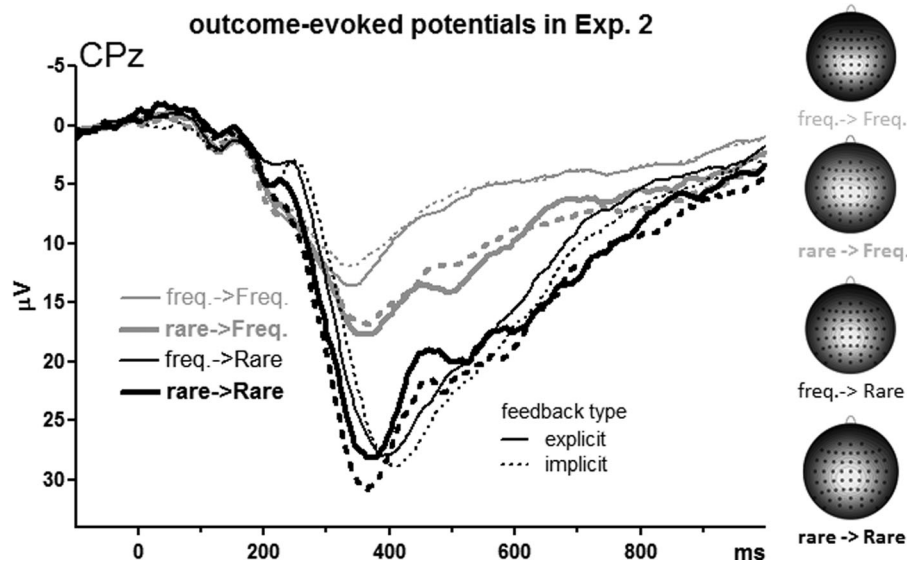
Figure 7. CPz waveforms and scalp topographies of P3 for each guess-outcome combination in Experiment 2. Waveforms are displayed in each of the four combinations: f → F (thin gray), r → F (bold gray), f → R (thin black), and r → R (bold black), separately for explicit (solid) and implicit (dotted) information about whether the guess was correct. Unit of *x* axis is ms, with the zero point denoting outcome onset. Unit of *y* axis is µV, with negative polarity plotted upwards. The maps display scalp topographies at the time points of the P3 peaks (340 ms, 370 ms, 390 ms, 380 ms for f → F, r → F, f → R, r → R respectively, pooled across the explicit and implicit conditions). View on the head is from above. Each map is scaled from minimum (black) to maximum (white).

were presented in black throughout. Thus, P3 evoked by these stimuli contained information about the preceding guess.

Throughout, independently from the preceding guess, P3s evoked by frequent stimuli were larger with explicit color-provided feedback than with implicit feedback. It must remain unclear whether this effect was due to explicitness of feedback or due to the physical saliency of color. Data from a recent study speak for the latter alternative. In that study (Verleger, Baur, Metzner, & Śmigasiewicz, 2014), color and letter stimuli were used (though not as feedback stimuli) and P3 amplitudes were globally increased when color was the relevant feature compared to when letter identity was the relevant feature. In the present study, this global increase of P3 by explicit color feedback did not occur for rare outcomes. Probably, the rare letter identity was so salient that the additional variation of color was less intrusive.

General Discussion

Previous evidence from prediction tasks, when either the next stimulus or its reward value had to be guessed, is inconsistent about whether P3 amplitudes evoked by outcome stimuli are larger after incorrect predictions (Hajcak et al., 2005, 2007; Sutton et al., 1965), or after correct predictions (Bellebaum et al., 2010; Bismark et al., 2013; Gentsch et al., 2013; Kimura & Katayama, 2013; Levit et al., 1973; Wu & Zhou, 2009; Yang et al., 2013), or do not differ between correct and incorrect predictions (Goyer et al., 2008; Kotchoubey et al., 1997; Verleger & Cohen, 1978; Yeung & Sanfey, 2004). It was conjectured by Sutton et al. (1965) that the effects of guess outcome on P3 depend on interactions with other factors, one of them being the ratio of occurrence of the two outcomes (Tueting et al., 1970). In the present study, we followed this line. Indeed, a consistent effect of prediction on P3 amplitudes was found for the frequent stimuli: P3 amplitudes were increased when frequent stimuli were incorrectly predicted. A smaller and inverted

effect of prediction on P3 amplitudes was found for the rare stimuli: P3 amplitudes were increased when rare stimuli were correctly predicted. These seemingly contradictory results may be reduced to one regularity: Outcome-evoked P3 amplitudes were increased when participants had chosen the rare prediction (leading to incorrectly predicted frequent stimuli and to correctly predicted rare stimuli).

We expected additive effects of guess probability and stimulus probability on P3 amplitudes. However, these two factors interacted: Making the rare prediction had large effects on the small P3 evoked by frequent stimuli and small effects on the large P3 evoked by rare stimuli. This pattern of results appears like a ceiling effect. Interestingly, it formally resembles the pattern reported by Munson et al. (1984) who analyzed effects of predictions on outcome-evoked P3 depending on which one of two equiprobable stimuli had preceded. In that study, repetition of preceding stimuli evoked a small P3, and alternations evoked a large P3. Formally similar to our result, the small P3 evoked by repeating stimuli was much increased when these stimuli were incorrectly predicted (as with our frequent stimuli) and the large P3 evoked by stimulus change was only little affected (as with our rare stimuli). Again, this appears like a ceiling effect. A nearby interpretation of these ceiling effects is that the process reflected by P3 was elicited in any case by rare stimuli (in our study) or stimulus alternations (in Munson et al.'s study) so could not be additionally increased by variations of the preceding guesses.

It may be suspected that the effects of guess frequency were mediated by effects of payoff because frequency (80%/20%) was confounded with payoff (2 ct./8 ct.). Increased payoff has been found to increase P3 amplitudes in gambling tasks (Bellebaum et al., 2010; Goyer et al., 2008; Yang et al., 2013; Yeung & Sanfey, 2004) though possibly moderated by other factors (Wu & Zhou, 2009). This suggestion leads to complex considerations, already for the apparently uncontroversial effect of stimulus probability (larger

P3 after rare than frequent stimuli, R > F). At first sight, this R > F effect may as well be due to, or amplified by, the larger payoff with rare than frequent stimuli (8 ct. vs. 2 ct.). But, this argument can only be made for correctly predicted stimuli. Because there was no gain with incorrectly guessed stimuli, nor was there a fine, the payoff was 0 ct. both for rare and frequent stimuli when incorrectly predicted, whereas the R > F amplitude effect also held for incorrectly predicted stimuli. Thus, the main effect of stimulus probability cannot be easily explained by the payoff variation between stimuli. However, one may argue that what matters is not the payoff associated with the stimuli but rather the discrepancy between expected payoff and actual payoff. Thus, the payoff argument would agree to the general idea that stimulus-evoked P3 reflects traces of the guessing behavior. Several formulas might be proposed how participants estimate the expected payoff and, consecutively, the discrepancy between actual and expected payoff, taking into account how these expectancies are weighted by event probabilities. The bottom line of this is that effects of payoff can only be excluded by conducting new experiments where other payoff arrangements are used or no payoff at all is associated to the stimuli. We may note, though, that Tueting et al. (1970) obtained effects of guess probability on P3 amplitudes similar to the present ones, without any payoff associated to successful guessing.

Therefore, we keep here to the notion that stimulus probability and guessing probability both have effects on P3 amplitude, with a ceiling effect for effects of guessing probability for rare stimuli, presumably because the process reflected by P3 is evoked by rare stimuli in any case. What can be the process that is reflected by outcome-evoked P3? Recently, we have stated the hypothesis that P3b reflects reactivation of inactive established stimulus-response (S-R) links (Verleger, Baur et al., 2014; Verleger, Metzner, Ouyang, Śmigasiewicz, & Zhou 2014; Verleger, Hamann, Asanowicz, & Śmigasiewicz, 2015).[1] This hypothesis was derived from choice-response tasks, where such S-R links might be "if X, press left," "if U, press right." When X is frequent and U is rare, X-left will be active most of the time, so U-right will have to be reactivated when U is perceived, and this activation is the process reflected by P3b. Applying this hypothesis to the guessing task meets with the problem that outcome stimuli in guessing tasks do not require any action, unlike the stimuli in choice-response tasks. Indeed, actions do not seem essential for these P3s to be elicited: Results did not dramatically differ from previously reported ones when participants were required to respond to the unpredictable stimuli, by pressing once more the associated key (Verleger & Cohen, 1978). We have recently confirmed this result in a comparison within subjects (unpublished data).

Can this S-R link hypothesis nevertheless be applied to guessing tasks? Why are large P3 amplitudes evoked by outcome stimuli even though no responses are required to these outcomes? It may be argued that responses are actually required to the outcome stimuli when participants play the guessing-task game, namely, the internal responses "I was right" or "I was wrong." These responses being internal ones ("right" and "wrong") does not seem to be a principal difference from oddball tasks where internal counting of the rare oddball stimuli has often been required as response. The effect of the preceding guess on these internal responses may be due to several factors. One possibility is that the frequent key press is not made for guessing but just for continuing the task such that participants are less interested in what the stimulus will actually be. In contrast, when making the rare guess, participants might be much more interested in the outcome, and it might be this increased level of interest that is responsible for larger P3 amplitudes. In terms of S-R link hypothesis, the internal response right or wrong might not get much activated after the more automatic frequent guess. What speaks against this hypothesis is that there was little difference between fixed and alternating key mapping in Experiment 1, although alternating key mapping reduces the automaticity of making frequent guesses to a large extent.

An alternative possibility is that the stimuli evoking these internal responses are not the isolated outcomes but the combinations of guesses and outcomes, to form event compounds.[2] Thus, the S-R links may be defined as (f → F, right), (r → F, wrong), (f → R, wrong), (r → R, right). By being evoked most of the time, (f → F, right) would need least reactivating and, thus, elicit the smallest P3. By this account, the prediction task differs from choice-response tasks by the fact that responses are required to combinations of two successive events rather than to one event. We are currently pursuing the question whether P3s evoked by the outcomes in prediction tasks follow the same regularities as P3s evoked by the second stimuli in choice response tasks in which participants are to choose their responses by combining two successive events.

There have not been many attempts to understand the processing reflected by outcome-evoked P3 in guessing tasks and to integrate such understanding to what is known about P3 from other tasks, most notably the oddball task. In view of a wealth of new findings, Sutton (1979) argued against the use of unitary concepts like the concept of information transmission favored by him before. Of interest, Donchin's (1981) context updating hypothesis, primarily derived from oddball tasks, was used to make the prediction in guessing tasks that P3 evoked by outcome stimuli would be related to participants' choices in the trials following these outcomes. Some effects were indeed found, although dependent on the preceding outcomes (Munson et al., 1984) but this line of research has not been pursued further. The present study aims at reviving such attempts of a deeper understanding of what is the common nature of the processes reflected by P3 in guessing tasks and other tasks.

---

1. This hypothesis is compatible with conclusions drawn by Verleger (1997) from his review on experimental effects on P3 latency and with the major idea put forward in Verleger et al. (2005) that P3b forms a bridge between S and R. It is certainly not compatible with Verleger's (1988) closure hypothesis, which states that P3b is evoked by expected rather than by unexpected events. This early hypothesis had been explicitly withdrawn by Verleger (1998; cf. Verleger et al., 2005), mainly because the closure idea neglected P3b's close association to response selection (cf. Verleger, 1997) and because its main prediction, that P3b is evoked by expected rather than by unexpected events, conflicted with several pieces of evidence, as convincingly summarized by Sommer, Leuthold, and Matt (1998).

---

2. Thanks to Istvan Czigler who made this suggestion about event compounds in an informal communication (Dortmund, May 7, 2014, conference of the European Society for Cognitive and Affective Neuroscience).

References

Bechara, A., Damasio, A. R., Damasio, H., & Anderson, S. W. (1994). Insensitivity to future consequences following damage to human prefrontal cortex. *Cognition*, *50*, 7–15.

Bellebaum, C., Polezzi, D., & Daum, I. (2010). It is less than you expected: The feedback-related negativity reflects violations of reward magnitude expectations. *Neuropsychologia*, *48*, 3343–3350. doi: 10.1016/j.neuropsychologia.2010.07.023

Bismark, A. W., Hajcak, G., Whitworth, N. M., & Allen, J. J. B. (2013). The role of outcome expectations in the generation of the feedback-related negativity. *Psychophysiology*, *50*, 125–133. doi: 10.1111/j.1469-8986.2012.01490.x

Davis, H., Mast, T., Yoshie, N., & Zerlin, S. (1966). The slow response of the human cortex to auditory stimuli: Recovery process. *Electroencephalography and Clinical Neurophysiology*, *21*, 105–113.

Debener, S., Makeig, S., Delorme, A., & Engel, A. K. (2005). What is novel in the novelty oddball paradigm? Functional significance of the novelty P3 event-related potential as revealed by independent component analysis. *Cognitive Brain Research*, *22*, 309–321.

Dien, J., Spencer, K. M., & Donchin, E. (2004). Parsing the late positive complex: Mental chronometry and the ERP components that inhabit the neighborhood of the P300. *Psychophysiology*, *41*, 665–678.

Donchin, E. (1981). Surprise! . Surprise? *Psychophysiology*, *18*, 493–513.

Donchin, E., & Cohen, L. (1969). Anticipation of relevant stimuli and evoked potentials: A reply *to* Näätänen. *Perceptual and Motor Skills*, *29*, 115–117.

Duncan-Johnson, C. C., & Donchin, E. (1977). On quantifying surprise: The variation of event-related potentials with subjective probability. *Psychophysiology*, *14*, 456–467.

Gaeta, H., Friedman, D., & Hunt, G. (2003). Stimulus characteristics and task category dissociate the anterior and posterior aspects of novelty P3. *Psychophysiology*, *40*, 198–208.

Gentsch, K., Grandjean, D., & Scherer, K. R. (2013). Temporal dynamics of event-related potentials related to goal conduciveness and power appraisals. *Psychophysiology*, *50*, 1010–1022. doi: 10.1111/psyp.12079

Goyer, J. P., Woldorff, M. G., & Huettel, S. A. (2008). Rapid electrophysiological brain responses are influenced by both valence and magnitude of monetary rewards. *Journal of Cognitive Neuroscience*, *20*, 2058–2069. doi: 10.1162/jocn.2008.20134

Hajcak, G., Holroyd, C. B., Moser, J. S., & Simons, R. F. (2005). Brain potentials associated with expected and unexpected good and bad outcomes. *Psychophysiology*, *42*, 161–170.

Hajcak, G., Moser, J. S., Holroyd, C. B., & Simons, R. F. (2007). It's worse than you thought: The feedback negativity and violations of reward prediction in gambling tasks. *Psychophysiology*, *44*, 905–912.

Kimura, K., & Katayama, J. (2013). Outcome evaluation in group decision making using the majority rule: An electrophysiological study. *Psychophysiology*, *50*, 848–857. doi: 10.1111/psyp.12068

Kotchoubey, B., End, H., Grözinger, B., Kornhuber, A., Jordan, J. S., & Kornhuber, H. H. (1994). Pre- and post-movement potentials in an informed guessing situation. *Journal of Psychophysiology*, *8*, 142–156.

Kotchoubey, B., Grözinger, B., Kornhuber, A. W., & Kornhuber, H. H. (1997). Electrophysiological analysis of expectancy: P3 in informed guessing. *International Journal of Neuroscience*, *91*, 105–122.

Levit, R. A., Sutton, S., & Zubin, J. (1973). Evoked potential correlates of information processing in psychiatric patients. *Psychological Medicine*, *3*, 487–494.

Munson, R., Ruchkin, D. S., Ritter, W., Sutton, S., & Squires, N. K. (1984). The relation of P3b to prior events and future behavior. *Biological Psychology*, *19*, 1–29.

Ritter, W., Vaughan H. G., Jr., & Costa, L. D. (1968). Orienting and habituation to auditory stimuli: A study of short term changes in averaged evoked responses. *Electroencephalography and Clinical Neurophysiology*, *25*, 550–556.

Simons, R. F., Öhman, A., & Lang, P. J. (1979). Anticipation and response set: Cortical, cardiac and electrodermal correlates. *Psychophysiology*, *16*, 222–233.

Sommer, W., Leuthold, H., & Matt, J. (1998). The expectancies that govern the P300 amplitude are mostly automatic and unconscious. (Continuing commentary). *Behavioral and Brain Sciences*, *21*, 149–150.

Squires, N. K., Squires, K. C., & Hillyard, S. A. (1975). Two varieties of long-latency positive waves evoked by unpredictable auditory stimuli in man. *Electroencephalography and Clinical Neurophysiology*, *38*, 387–401.

Sutton, S. (1979). P300—Thirteen years later. In H. Begleiter (Ed.), *Evoked brain potentials and behavior* (pp. 107–126). New York, NY: Plenum Press.

Sutton, S., Braren, M., Zubin, J., & John, E. R. (1965). Evoked-potential correlates of stimulus uncertainty. *Science*, *150*, 1187–1188.

Tueting, P., Sutton, S., & Zubin, J. (1970). Quantitative evoked potential correlates of the probability of events. *Psychophysiology*, *7*, 385–394.

Urbach, T. P., & Kutas, M. (2006). Interpreting event-related brain potential (ERP) distributions: Implications of baseline potentials and variability with application to amplitude normalization by vector scaling. *Biological Psychology*, *72*, 333–343.

Verleger, R. (1988). Event-related potentials and cognition: A critique of the context updating hypothesis and an alternative interpretation of P3. *Behavioral and Brain Sciences*, *11*, 343–356.

Verleger, R. (1997). On the utility of P3 latency as an index of mental chronometry. *Psychophysiology*, *34*, 131–156.

Verleger, R. (1998). Towards an integration of P3 research with cognitive neuroscience. (Author's response on continuing commentary). *Behavioral and Brain Sciences*, *21*, 150–152.

Verleger, R. Baur, N., Metzner, M. F., & Śmigasiewicz, K. (2014). The hard oddball: Effects of difficult response selection on stimulus-related P3 and on response-related negative potentials. *Psychophysiology*, *51*, 1089–1100. doi: 10.1111/psyp.12262

Verleger, R., & Cohen, R. (1978). Effects of certainty, modality shift and guess outcome on evoked potentials and reaction times in chronic schizophrenics. *Psychological Medicine*, *8*, 81–93.

Verleger, R., Hamann, L. M., Asanowicz, D., & Śmigasiewicz, K. (2015). Testing the S-R link hypothesis of P3b: The oddball effect on S1-evoked P3 gets reduced by increased task relevance of S2. *Biological Psychology*. Advance online publication. doi: 10.1016/j.biopsycho.2015.02.010

Verleger, R., Jaśkowski, P., & Wascher, E. (2005). Evidence for an integrative role of P3b in linking reaction to perception. *Journal of Psychophysiology*, *19*, 165–181.

Verleger, R. Metzner, M. F., Ouyang, G., Śmigasiewicz, K., & Zhou, C. (2014). Testing the stimulus-to-response bridging function of the oddball-P3 by delayed response signals and residue iteration decomposition (RIDE). *NeuroImage*, *100*, 271–280. doi: 10.1016/j.neuroimage.2014.06.036

Walter, W. G., Cooper, R., Aldridge, V. J., McCallum, W. C., & Winter, A. L. (1964). Contingent negative variation: An electric sign of sensorimotor association and expectancy in the human brain. *Nature*, *203*, 380–384.

Wu, Y., & Zhou, X. (2009). The P300 and reward valence, magnitude, and expectancy in outcome evaluation. *Brain Research*, *1286*, 114–122. doi: 10.1016/j.brainres.2009.06.032

Yang, Q., Gu, R., Tang, P., & Luo, Y.-J. (2013). How does cognitive reappraisal affect the response to gains and losses? *Psychophysiology*, *50*, 1094–1103. doi: 10.1111/psyp.12091

Yeung, N., & Sanfey, A. G. (2004). Independent coding of reward magnitude and valence in the human brain. *Journal of Neuroscience*, *24*, 6258–6264.