Error-likelihood prediction in the medial frontal cortex:

A critical evaluation

Sander Nieuwenhuis ¹,²,³, Tanja Sophie Schweizer ³,
Rogier B. Mars ⁴,⁵, Matthew M. Botvinick ⁶, & Greg Hajcak ⁷

¹ Dept of Psychology, Leiden University, Leiden, the Netherlands
² Leiden Institute for Brain and Cognition, Leiden University, Leiden, the Netherlands
³ Dept of Psychology, Vrije Universiteit, Amsterdam, the Netherlands
⁴ F.C. Donders Centre for Cognitive Neuroimaging, Nijmegen, The Netherlands
⁵ Nijmegen Institute for Cognition and Information,
   Radboud Universiteit Nijmegen, The Netherlands
⁶ Dept of Psychology, Dept of Psychiatry, and Center for Cognitive Neuroscience,
   University of Pennsylvania, Philadelphia, PA, USA
⁷ Dept of Psychology, Stony Brook University, Stony Brook, NY, USA

Address correspondence to: Sander Nieuwenhuis, Leiden University, Department of Psychology,
Cognitive Psychology Unit, Wassenaarseweg 52, 2333 AK Leiden, The Netherlands. Phone: +31-71-5273834. Fax: +31-71-5273783. E-mail: snieuwenhuis@fsw.leidenuniv.nl.
Abstract

A recent study has proposed that posterior regions of the medial frontal cortex (pMFC) learn to predict the likelihood of errors occurring in a given task context. A key prediction of the error-likeness hypothesis is that the pMFC should exhibit enhanced activity to cues that are predictive of high compared to low error rates. We conducted three experiments, two using functional neuroimaging and one using event-related potentials, to test this prediction in human volunteers. The three experiments replicated previous research in showing clear evidence of increased pMFC activity associated with errors, conflict, negative feedback, and other aspects of task performance. However, none of the experiments yielded evidence for an effect of cue signalled error likelihood on pMFC activity, or any indication that such an effect developed with learning. We conclude that although the error-likeness hypothesis presents an elegant integrative account of medial frontal cortex function, it requires additional empirical support to remain tenable.

Keywords: ACC, dopamine, error processing, fMRI, reinforcement learning, stop signal
During the past decade there has been a surge of interest in the functions of the anterior cingulate cortex (ACC) and adjoining, more dorsal areas of the medial frontal cortex. This region is activated by a broad range of demanding cognitive tasks, as indicated by numerous functional neuroimaging studies (Duncan and Owen 2000). However, despite the presence of a rich empirical database, there is still little consensus regarding the specific role of the medial frontal cortex in cognitive function: Neurophysiological studies in animals and human electrophysiological and neuroimaging studies have implicated parts of the medial frontal cortex in the detection of errors (Falkenstein and others 2000; Ito and others 2003), monitoring of response conflict (Botvinick and others 2004), goal-based action selection (Matsumoto and Tanaka 2004), reinforcement learning (Holroyd and Coles 2002), and other evaluative functions (for reviews, see Ridderinkhof and others 2004; Rushworth and others 2004). Yet, despite the diversity of these proposed functions, the critical areas identified by these studies all largely fall within the same posterior region of the medial frontal cortex (pMFC), a region that encompasses large parts of the caudal anterior cingulate cortex and the pre-supplementary motor area (pre-SMA)(Picard and Strick 1996; Ridderinkhof and others 2004). Indeed, some of the mentioned functions seem to be carried out by separate yet closely intermingled cells or cell populations in the pMFC (e.g., Ito and others 2003),

Given this pattern of results, it seems plausible that some of the proposed pMFC functions are in fact constituents of one superordinate function. According to this view, the challenge is to develop an umbrella theory that subsumes several current theories of pMFC function. Recently, an apparent step in this direction has been taken by Brown and Braver (2005a). These authors proposed an elegant new hypothesis of pMFC function according to which the pMFC codes the predicted likelihood of errors occurring in a specific task context. Brown and Braver propose that through experience, the pMFC gradually learns to associate task contexts with error likelihood, on the basis of dopaminergic reinforcement learning signals. Thus, when a particular task context is
encountered, neurons in the pMFC increase their activity by an amount proportional to the likelihood of errors—a signal that may serve as an early warning for the cognitive system. According to this error-likelihood hypothesis, errors and response conflict activate the pMFC because both are circumstances that predict undesired consequences. The hypothesis also incorporates many aspects of a previous theory that emphasizes the interaction between the pMFC and the dopamine system in reinforcement learning (Holroyd and Coles 2002), although according to that theory the association between task contexts and outcomes occurs in the basal ganglia instead of the pMFC. The experiments described in the present paper were conducted to provide a critical test of the error-likelihood hypothesis.

A key prediction of the error-likelihood hypothesis that distinguishes it from previously proposed theories is that the pMFC can learn to associate arbitrary stimulus features with error likelihood. Furthermore, the error-likelihood hypothesis holds that learning these correlations between stimulus features and trial outcomes can occur on the basis of relatively little experience, and takes place irrespective of whether or not a stimulus feature is associated with a response or perceived by a participant as task-relevant. Previous theories that have implicated the pMFC in reinforcement learning have specifically stressed its role in encoding the relationship between actions and the value of their outcomes (Holroyd and Coles 2002; Rushworth and others 2004), suggesting that stimulus-reward associations are formed in other brain areas (Pears and others 2003).

Brown and Braver (2005a) tested the predictions of the error-likelihood hypothesis using a stop-change task (Logan and Burkell 1986) in combination with functional magnetic resonance imaging (fMRI). Participants were required to produce speeded, spatially compatible button-press responses to left- and right-pointing arrows. On one-third of the trials, the first arrow was rapidly followed by a second arrow pointing in the opposite direction, requiring participants to stop their ongoing response and produce the opposite response instead. The second arrow was presented at
variable time intervals after the first, calibrated so as to produce error rates (i.e., percentage unsuccessfully reversed responses) of 50% and 4% in the high and low error-likelihood conditions, respectively. Importantly, the error likelihood on each trial (i.e., high or low) was indicated by a color cue that preceded the onset of the first arrow stimulus. Participants were not informed of the meaning of the cue color. Nevertheless, the fMRI analyses revealed two areas in the pMFC that showed increased activity on high compared to low error-likelihood trials, even when only correct trials were considered. Furthermore, for one of these areas, in the pre-SMA, the error-likelihood effect gradually emerged over the course of the experiment\(^1\). The authors concluded that their results are consistent with a role for the pMFC in learning to predict error likelihood.

We conducted three experiments, two using fMRI and one using event-related potentials (ERPs), in an attempt to critically evaluate Brown and Braver’s (2005a) findings. In Experiment 1 we investigated whether the pMFC would show similar sensitivity to error likelihood in an entirely different experimental paradigm (visual search). As discussed below, the original purpose of this fMRI experiment was to test an alternative hypothesis of the effect reported by Brown and Braver. The results of this experiment led us to conduct Experiments 2 (fMRI) and 3 (ERP), in which we attempted to replicate Brown and Braver’s results using their own stop-change task, while trying to avoid some methodological limitations of the original study.

To look ahead briefly, we did not find any evidence for the prediction of the error-likelihood hypothesis that pMFC activity is modulated by cue-induced error-likelihood predictions. However, consistent with previous research, all three experiments provided clear evidence of increased pMFC activity associated with negative feedback, errors, conflict, and demanding cognitive tasks (Botvinick and others 2004; Ridderinkhof and others 2004), results that have been taken as support for competing accounts of pMFC function. Together, these results cast serious doubt on the error-likelihood hypothesis.
Experiment 1

In Experiment 1 we tested the error-likelihood hypothesis using a visual search task. On each trial participants searched for a target among multiple distractors, responded according to the identity of the target, and then received performance feedback. To ensure that participants made errors, the search display was masked after a brief individually-calibrated duration. On half of the trials the target was defined by a unique feature (i.e., pop-out search), resulting in a rapid search process with relatively few errors (low error-likelihood condition). On the other half of the trials the target was defined by a conjunction of features (conjunction search), resulting in a slow search process with many errors (high error-likelihood condition). As in Brown and Braver (2005a), an arbitrary visual cue, presented before the onset of the search display, signaled whether the trial was associated with a high or low error likelihood.

The critical prediction, based on the error-likelihood hypothesis and Brown and Braver’s (2005a) fMRI results, was that the pMFC should be more active following the high error-likelihood cue than following the low error-likelihood cue. In addition to testing this prediction, we also examined the effects of two variables that have repeatedly been shown to modulate pMFC activity: task difficulty (conjunction search vs. pop-out search; Paus and others 1998) and feedback valence (positive vs. negative; Holroyd and others 2004b).

It is important to note that differences in error rate are typically confounded with differences in task difficulty (i.e., the number and complexity of the required mental computations). That is, one generally makes more errors when a task is more difficult. Accordingly, our experiment and that of Brown and Braver (2005a) cannot exclude the possibility that the pMFC anticipates task difficulty or the amount of required mental effort rather than the likelihood of errors. However, the current task was designed such that it could easily be modified to test this alternative hypothesis regarding the nature of pMFC representations. Specifically, by taking out the masks and requiring participants to make speeded responses (while keeping the number of errors to a minimum), the
cues would signal two conditions that differed in task difficulty (as evidenced by increased conjunction-search reaction times) but presumably not in error rate. This idea formed the basis for a potential follow-up experiment, planned in case the key prediction of the current experiment was confirmed.

Materials and Methods

Participants. Participants were 14 young adults (10 female; average age 22.3 years). All participants were right-handed and all had normal or corrected-to-normal visual acuity. They were paid €15 for a 1.5-h session. Written informed consent was obtained from all participants in all three reported experiments, and the experiments were approved by the research ethics committee of the Vrije Universiteit Medical Center.

Task. Each trial started with the presentation of an error-likelihood cue: a schematic outline of a lightbulb or a bicycle, presented in the center of the screen for a random real-number interval between 1.0 and 6.0 s (Figure 1). On each trial, cue identity (lightbulb or bicycle) predicted whether the trial would be associated with low or high error likelihood. The two cues were equiprobable and presented in a quasi-random order, and the mapping of cue identity to high vs. low error likelihood was counterbalanced across participants. Immediately following cue offset, a search display consisting of one target item (a red tilted bar) and 16 distractor items (described further below) was presented. Participants were instructed to find the target item and to press a button with their right index finger if the bar was tilted to the left, or with their right middle finger if the bar was tilted to the right. After a variable duration, each item in the search display was replaced with a pattern mask so that the target item could no longer be discriminated. Participants were encouraged to give a response even if they had not found the target item in time. The masked search display remained visible until a response deadline of 2000 ms after the onset of the search display. During the next 500 ms, a feedback stimulus was presented that evaluated the response: “correct” (green font),
“error” (red font), or “too late!!” (yellow font). Subsequently the screen was blank for a random intertrial interval between 1.5 and 15.0 s, after which the next trial started. The interval between the error-likelihood cue and the search display as well as the intertrial interval were jittered in order to de-correlate the hemodynamic signals associated with the cue and other stimulus and response events (Burock and others 1998).

On low error-likelihood trials, the distractors in the search display were green tilted bars. Because in this case the target is defined by a unique feature (color), it “pops out” from the display, resulting in a fast and efficient search process (pop-out search). On high error-likelihood trials, the distractors were green tilted bars and red vertical bars. In this case the target does not consist of a single identifying feature but is defined by a specific conjunction of features (color and orientation), resulting in a more time-consuming and error-prone search process (conjunction search). The error rate in the high error-likelihood condition was controlled by dynamically adjusting the presentation duration of the search display by means of a staircase tracking algorithm. The presentation duration of the search display was incremented by 80 ms for each incorrect conjunction-search trial and decremented by 40 ms for each correct conjunction-search trial. Presentation duration was initialized at 600 ms. The algorithm aimed at 33% errors in the high error-likelihood condition.

Participants received instructions and 20 practice trials outside the scanner before entering the experimental phase. The experimental phase consisted of 180 trials, divided into six equal blocks, with short breaks in between. The task instructions encouraged participants to respond as quickly as possible while minimizing the number of errors. Participants were also instructed to “pay close attention to the pictures of the lightbulb and bicycle, because they provide information about the task. The lightbulb always precedes one type of trial, and the bicycle always precedes another type of trial”. This instruction was accompanied by an illustration of the sequence of events on a pop-out search trial and on a conjunction search trial. The goal of this instruction was to promote explicit learning of the relationship between the cues and error likelihood, and thereby to increase
chances of observing differential neural activity related to the predictive nature of the two cues. An informal exit interview revealed that during the experiment 10 out of 14 participants were aware of the relationship between cue identity and task difficulty.

**Stimuli.** Stimuli (see Figure 1) were presented in color against a black visual display projected into the scanner. The cues consisted of a white outline of a lightbulb or a bicycle and subtended approximately 5.7°. The search display consisted of 17 items that were randomly plotted in the cells of an imaginary 5 x 5 matrix (12.6° x 12.6°), with some random jitter within the cells. The target item was a red bar (1.4° x 0.14°), tilted 45° to the left or to the right with each orientation occurring equally often in each block. On pop-out search trials (i.e., low error likelihood), the distractor items were 16 green bars (1.4° x 0.14°), tilted 45° to the left or to the right, the orientation determined randomly for each item. On conjunction search trials (high error likelihood), the distractor items were 8 green tilted bars and 8 red vertical bars. In the mask display each of the search items was masked by adding three tilted, red- or green-colored bars.

**fMRI image acquisition.** Images were collected with a 1.5-T Siemens Sonata scanner equipped with a volume head coil. Anatomical images were collected using a T1-weighted MP-RAGE sequence (TR=2700 ms, TE=3.95 ms, TI=950 ms, FA=8°, 256*160 coronal matrix, 1.0*1.0 mm in-plane resolution, 224 1.1-mm slices). Functional images were reconstructed from twenty oblique slices acquired using a T2*-weighted EPI sequence (TR=2000 ms, TE=60 ms, FA=90°, 64*64 matrix, 3.0*3.0 mm in-plane resolution, 5.0-mm slices, 20% gap). Image acquisition varied across trials with respect to stimulus onset, yielding an effectively higher temporal sampling rate (Miezin and others 2000). Six functional runs (228 scans each) were collected. The first two scans of each run were discarded because they were recorded before the longitudinal magnetization reached a steady state recovery value.

**fMRI image analysis.** Data were preprocessed and analyzed with BrainVoyager software (Maastricht, The Netherlands). Image preprocessing consisted of: rigid-body 3D motion correction
using trilinear interpolation; slice scan time correction using sinc interpolation; spatial smoothing with a 4-mm fullwidth at half maximum (FWHM) Gaussian kernel; voxel-wise linear detrending; high-pass filtering (above 7 cycles per time course) to remove low frequencies; and low-pass filtering with a 2.8-sec FWHM Gaussian kernel to remove high frequencies. Spatial normalization was performed using the standard 9-parameter landmark method of Talairach and Tournoux (1988). Images were resampled into 1-mm cubic voxels using sinc interpolation. For each participant, the blood oxygen-level dependent (BOLD) responses across the scanning run were modelled with a general linear model that included seven regressors. Two regressors accounted for the high and low error-likelihood cues. Two additional regressors modelled the pop-out search display and the conjunction-search display, including the response to these displays. Finally, three regressors accounted for the three possible feedback stimuli (i.e., correct, incorrect, too late). Correlations between the cue predictors and the other predictors were all below .30. The hemodynamic response to each event was estimated by convolving each regressor with a standard gamma function (Boynton and others 1996). For each voxel and each event type, a parameter estimate was generated that indicated the strength of covariance between the data and the hemodynamic response function; these estimates were corrected for temporal autocorrelation using a first-order autoregressive model. Contrasts between parameter estimates for different events were calculated for each participant, and the results submitted to a group analysis that treated inter-subject variability as a random effect. Statistical parametric maps were derived from the resulting t-values associated with each voxel and were thresholded at a conservative value (p < .0005, uncorrected), with a contiguity threshold of 120 mm³ as a further precaution against type-1 errors (Forman and others 1995). In case this whole-brain analysis did not reveal any clusters of activation in the pMFC, it was followed by a region-of-interest (ROI) analysis focusing on the pMFC and using a more liberal threshold (p < .005, uncorrected). The pMFC ROI was loosely defined as the area in the medial wall superior to the corpus callosum, posterior to the genu, and anterior to Y=0 mm. As
discussed above, this area is consistently activated in studies of performance monitoring (Ridderinkhof and others 2004). The location of the peak activity associated with each cluster of activation was reported in Talairach coordinates (Talairach and Tournoux 1988).

Results

**Behavior.** The average error rates on conjunction-search (high error likelihood) and pop-out search trials (low error likelihood) were 29.4% and 2.7%, respectively, $F(1,13) = 1807.4, p < .001$. The corresponding average correct reaction times (RTs) were 1019 ms and 791 ms, $F(1,13) = 72.5, p < .001$. The mean presentation duration of the search display was 356 ms.

**fMRI.** To identify brain areas that were sensitive to error-likelihood signals, we performed the following contrast: high > low error-likelihood cue. A whole-brain analysis indicated that there were no brain areas that exhibited greater activity for the high vs. the low error-likelihood cue. A subsequent ROI analysis focusing on the pMFC also revealed no differential activity associated with the two cues. This was not due to a lack of power; a whole-brain exploratory analysis using the same liberal threshold ($p < .005$, uncorrected) revealed 14 areas outside the pMFC (including the bilateral striatum, parahippocampal gyrus, precentral gyrus, and bilateral fusiform gyrus) that were more active following the high error-likelihood cue. The same pattern of results was obtained when we restricted the analyses to the second half of the experiment (blocks 4-6), by which one might expect the cognitive system to have correctly acquired the error likelihoods associated with the two cues (Brown and Braver 2005a).

In a further attempt to reveal potential error-likelihood effects, we focused directly on the two pMFC foci reported by Brown and Braver (2005a), and on the pMFC region that was most reliably activated by the two cues in our experiment. This latter region was identified by a conjunction analysis focusing on the contrast ([high error-likelihood cue > baseline] AND [low error-likelihood cue > baseline]), which revealed a substantial activation cluster in the caudal
pMFC (peak coordinates X = -4, Y = 4, Z = 51). We defined 10-mm cubic areas centered at each of these three ROIs, and for these areas computed and statistically compared (with $\alpha = .05$) the average regression coefficients associated with the high and low error-likelihood cue regressors in the general linear model. The resulting averages (standard errors) for high vs. low error-likelihood cues were (i) Brown and Braver’s ACC ROI: 2.15 (0.95) vs. 0.73 (0.40), $t(13) = 1.19$, $p = 0.26$; (ii) Brown and Braver’s pre-SMA ROI: 1.26 (.79) vs. -0.10 (0.30), $t(13) = 1.72$, $p = 0.11$; and (iii) our cue-sensitive ROI: 3.45 (1.70) vs. 1.99 (se=1.18), $t(13) < 1$. Altogether, these results argue against the presence in our data of a significant effect of error likelihood in the pMFC.

In an additional analysis we contrasted conjunction search and pop-out search to identify brain areas that were sensitive to the increased mental effort and additional cognitive operations associated with conjunction search. Not surprisingly, conjunction search was associated with increased activity in widespread areas across the brain. Most notably, the analysis revealed a large cluster of activation extending from the ACC into the pre-SMA, and activation clusters in the bilateral insula, and left intraparietal sulcus (Table 1). In contrast, the posterior cingulate cortex and subgenual rostral ACC showed deactivation during conjunction search. The deactivation of these areas with demanding cognitive activity is a common finding (Gusnard and Raichle 2001).

Finally, we identified several brain areas that were differentially sensitive to negative and positive feedback (Table 1). Of most relevance for the present purposes, an ROI analysis revealed three pMFC activation clusters (1 in ACC, 2 in pre-SMA) that showed greater activity to negative vs. positive feedback. No pMFC regions showed the opposite pattern.

**Discussion**

In Experiment 1 we found no evidence in support of the error-likelihood hypothesis. Contrary to the prediction of this hypothesis, activity in the pMFC was not reliably influenced by the identity of the cue, which predicted error likelihood. An analysis restricted to the second half of
the experiment also found no indication that error-likelihood predictions in the pMFC emerged over the course of the experiment. In contrast, in line with previous studies (Paus and others 1998; Duncan and Owen 2000; Holroyd and others 2004b; Mars and others 2005; but see Nieuwenhuis and others 2005), pMFC activity was reliably increased in the most demanding task condition (conjunction search), and following negative performance feedback. This demonstrates that our experimental design and scanning parameters were sufficiently sensitive to detect changes in pMFC activity.

A possible discrepancy between our findings and those of Brown and Braver (2005a) is that the two studies involved different types of errors. In the stop-change task used by Brown and Braver, participants are usually aware of the error while they are making it; the change signal arrives just too late to instigate a timely reversal of the response, often leading to an immediate emotional response (Hajcak and others 2003). Most research on the role of pMFC in error monitoring has focused on action slips of this type. In contrast, in our visual search task most errors are due to data limitations (i.e., insufficient perceptual evidence): The participant fails to find the target in time, has to guess a response, and learns from the feedback stimulus whether the guess was right or wrong. It is possible that the cognitive system treats or values this type of errors in a different way than action slips. This raises the possibility that the pMFC may not learn to predict this type of errors, and hence, that the error-likelihood hypothesis may not apply in the current task context. To address this concern, we conducted a second experiment in which we used the stop-change paradigm employed by Brown and Braver.

**Experiment 2**

In Experiment 2 we again tested the prediction of the error-likelihood hypothesis that pMFC activity should be modulated by the error likelihood associated with different task cues. However, in this experiment we employed the original stop-change paradigm utilized by Brown and Braver.
If our failure in Experiment 1 to confirm the prediction of the error-likelihood hypothesis was due to differences in the employed task or type of errors, then we should be able to replicate Brown and Braver’s findings when using their task.

One essential procedural change that we made to the original task concerned the time interval between the cues and subsequent target stimuli: Whereas Brown and Braver’s (2005a) experiment employed a short and fixed interval (1 sec) between the onset of the cue and the target, the present experiment used a relatively long, and variable, cue-target interval. Although we defer a complete discussion of this design issue to the Discussion of Experiment 2, a longer and variable cue-target interval was utilized so that the hemodynamic responses associated with the error-likelihood cues and other task events (e.g., target processing and responding) could be separated.

Although Brown and Braver (2005a) do not explicitly mention this, their task design (i.e., a short interval between cues and go signals) leaves open the possibility that the error-likelihood effects obtained in their experiment were driven not by the error-likelihood cues but by the go signals and/or associated responses. This possibility would seem more consistent with previous theories of the role of the pMFC in reinforcement learning, which have stressed its importance in encoding the relationship between actions and outcomes (Holroyd and Coles 2002; Rushworth and others 2004). As noted above, the task design in Experiment 2 allowed us to compute distinct estimates of the BOLD responses associated with the target/response period. To test the possibility that an error-likelihood effect occurred during this period, we contrasted go trials following high vs. low error-likelihood cues. The similar RTs in these two conditions (see Results) suggested that this contrast was not confounded by differences in processing of and responding to the go signal. Therefore, the analysis should yield a relatively pure measure of error-likelihood modulations during the target/response period.

Materials and Methods
Participants. Participants were 14 young adults (8 female; average age 23.7 years). All participants were right-handed and all had normal or corrected-to-normal visual acuity. They were paid €15 for a 1.5-h session.

Task and Stimuli. Each trial started with the presentation of an error-likelihood cue: a white or blue horizontal bar, presented in the center of the screen for a random interval between 1.0 and 5.0 s. On each trial, cue color predicted whether the trial would be associated with a low or high error likelihood. The two colors were equiprobable and presented in quasi-random order, and the mapping of cue color to high vs. low error likelihood was counterbalanced across participants. Immediately following cue offset, a go signal was presented that indicated the required button-press response. The go signal consisted of a left- or right-pointing arrow (constructed by adding an arrow head to the cue) with the left-pointing arrow requiring a response with the right index finger, and the right-pointing arrow requiring a response with the right middle finger. On 33% of the trials, a change signal was added to the go signal after a variable delay hereafter referred to as the change signal delay (CSD). The change signal consisted of a second, larger arrow appearing above the first and pointing in the opposite direction. The change signal indicated that the response had to be left-right reversed from the response indicated by the go signal. All stimuli for a given trial were the same color. Both go and change signals remained visible until a response deadline of 1000 ms after go signal onset. Subsequently the screen was blank for a random intertrial interval between 1.5 and 15.0 s, after which the next trial started. The interval between the cue and the go signal, and the intertrial interval, were jittered in order to de-correlate the hemodynamic signals associated with the cue and other stimulus events (Burock and others 1998).

Stimuli were presented in white or blue (RGB 128, 255, 255) on a black background. The cues subtended 2.5° x 0.9°. The go signal subtended 3.4° horizontally. The change-signal arrow was exactly twice as large as the go-signal arrow, and the vertical distance between the two arrows subtended 2.6°.
Error rates were controlled by dynamically and independently adjusting the CSDs for each error-likelihood condition by means of a staircase tracking algorithm. CSDs were shorter in the low than in the high error-likelihood condition, reflecting the well-established positive monotonic relationship between CSD and error rate in stop-signal tasks (Logan 1994). The CSD for the high error-likelihood condition was incremented by 50 ms for each correct high-error/change trial, and decremented by 50 ms for each incorrect high-error/change trial, aiming at 50% errors in this condition. The CSD for the low error-likelihood condition was incremented by 5 ms for each correct low-error/change trial, and decremented by 50 ms for each incorrect low-error/change trial, aiming at 10% errors. The CSDs for the high and low error-likelihood conditions were initialized at 300 ms and 100 ms, respectively.

Participants received instructions and 20 practice trials outside the scanner before entering the experimental phase. The experimental phase consisted of 240 trials altogether (of which 80 were change trials, 40 in each error-likelihood condition), divided into six equal blocks, with short breaks in between. When a change signal occurred, participants were instructed to stop their initial response and to give the opposite response as quickly as possible. They were told that unsuccessful change trials (i.e., initial response not stopped in time) constituted errors, and that they should try to minimize the number of errors. Participants were strongly discouraged from delaying their response to the go signal in anticipation of a possible change signal. They were further told to pay close attention to the color of the cue, because this signalled the CSD. The relationship between CSD and task difficulty was explained. Participants were not informed beforehand that the error rates for both cues were controlled by the experiment software.

fMRI image acquisition and image analysis. All details were the same as in Experiment 1, except for the following. Six functional runs (246 scans each) were collected. For each participant, the BOLD responses across the scanning run were modelled with a general linear model that included eight regressors. Two regressors accounted for the high and low error-
likelihood cues. Six regressors modeled the possible conjunctions of target/response type (go, successful change, unsuccessful change) and error likelihood (high, low). The onset of these regressors was time-locked to the presentation of the go signal. Correlations for the various pairs of predictors were all below .30. Thus, whereas Brown and Braver (2005a) used a single regressor per trial to model the various stimuli and the response, our general linear model allowed us to obtain separate estimates of the BOLD responses associated with the cues and with other trial events.

Results

Behavior. The average error rates on high and low error-likelihood change trials were 52.0% and 18.0%, respectively, $F(1,13)=383.4$, $p<.001$. The corresponding average correct RTs on change trials were 455 ms and 440 ms, $F(1,13)=13.2$, $p=.003$. Errors on go trials were rare (~1%). The RTs on correct go trials associated with high (439 ms) and low error-likelihood cues (437 ms) were similar, $F<1$. The mean CSDs on high and low error-likelihood trials were 162 ms and 77 ms, respectively.

We also examined whether participants adjusted their behavior in response to various trial types. These sequential-effect analyses indicated that RTs on correct go trials were slower if the immediately preceding trial was a successful change trial (464 ms) as compared to a correct go trial (428 ms), $F(1,13)=25.3$, $p<.0001$. Go RT was not systematically modulated by the error likelihood of the previous successful change trial (high 461 ms, low 467 ms), $F(1,13)=1.8$, $p=.20$.

fMRI. A whole-brain analysis did not reveal any brain areas that showed differential sensitivity to the high and low error-likelihood cues. Moreover, a follow-up ROI analysis yielded no pMFC regions that were sensitive to error likelihood. As in Experiment 1, the same pattern of results was obtained when the analyses were restricted to the second half of the experiment (blocks 4-6), by which learning of error likelihoods had had more time to take place.
Despite the absence of a significant error-likelihood effect, a conjunction analysis contrasting each of the cues against baseline indicated that both of the cues were associated with reliable activation in the pMFC (peak coordinates X = 2, Y = 0, Z = 51). To further increase statistical sensitivity to potential error-likelihood effects, we focused on this cue-sensitive pMFC area and on the two pMFC foci reported by Brown and Braver (2005a), employing the same procedure as in Experiment 1. The average regression coefficients (standard errors) associated with high vs. low error-likelihood cues were (i) Brown and Braver’s ACC ROI: 3.01 (0.69) vs. 2.13 (0.67), \(t(13) = 1.50, p = 0.16\); (ii) Brown and Braver’s pre-SMA ROI: 1.55 (1.10) vs. 0.81 (0.51), \(t(13) < 1\); and (iii) our cue-sensitive ROI: 5.83 (1.11) vs. 5.35 (1.22), \(t(13) < 1\). Thus, as in Experiment 1, none of the specific regions of interest within the pMFC showed a reliable cue-related effect of error likelihood, even at a significance threshold of \(\alpha = .05\) (uncorrected).

To evaluate the possibility that an error-likelihood effect occurred during the target/response-portion of the trial (instead of during the cue-target interval), we compared the fMRI response to go signals preceded by high and low error-likelihood cues. However, neither a whole-brain analysis nor a pMFC ROI analysis revealed significant differences for this contrast. Furthermore, the average regression coefficients (standard errors) associated with high vs. low error-likelihood cues did not reliably differ (\(\alpha = 0.05\)) for the cue-sensitive pMFC area in the present experiment, or for either of the two pMFC foci reported by Brown and Braver (2005a): (i) Brown and Braver’s ACC ROI: 3.77 (0.66) vs. 3.66 (0.54), \(t(13) < 1\); (ii) Brown and Braver’s pre-SMA ROI: 1.83 (0.85) vs. 1.69 (0.51), \(t(13) < 1\); and (iii) our cue-sensitive ROI: 6.19 (1.13) vs. 4.81 (1.08), \(t(13) = 1.84, p = 0.09\).

To identify brain areas that were sensitive to the processing conflict induced by change signals, we contrasted successful change trials vs. go trials. A whole-brain analysis revealed a region in the pre-SMA that was more active on change trials than on go trials. In addition, the analysis revealed increased, change-related activity in the right inferior frontal gyrus, an area
commonly associated with the suppression of ongoing responses (Aron and others 2003), the left inferior frontal gyrus, bilateral insula, and the precentral gyrus (Table 2).

Finally, we compared unsuccessful and successful change trials to reveal brain areas sensitive to errors. This contrast revealed highly reliable error-related activity in the ACC, and in a region of the medial frontal gyrus (Brodmann area 8, near the border with area 32) that is part of our pMFC ROI (see Ridderinkhof and others 2004). The right insula and left inferior parietal lobule also showed significant error-related activity (Table 2).

**Discussion**

Although in Experiment 2 we used the same experimental paradigm as Brown and Braver (2005a), we did not find support for the error-likelihood hypothesis, which proposes that the pMFC learns to predict the error likelihood associated with a given context. In particular, we did not manage to replicate Brown and Braver’s finding that the pMFC exhibits differential responses to cues predicting high and low error likelihood. We also did not find evidence that such differential responses occurred in conjunction with the go signal and associated response. As in Experiment 1, these negative results stood in marked contrast to the clear pMFC effects observed in response to other task variables: Increased response conflict, as associated with successful change trials compared to go trials, was characterized by substantial modulation of pre-SMA activity. Furthermore, errors following a change signal were associated with increased activity in the ACC. These findings are consistent with a large number of previous studies that have reported sensitivity of the pMFC to response conflict and errors (Botvinick and others 2004; Ridderinkhof and others 2004).

How can we explain the discrepancy between the current results and those of Brown and Braver (2005a)? As described above, an important difference between the two studies concerned the duration of the cue-target interval. In Brown and Braver’s study this interval had a short and
fixed duration, precluding the accurate calculation of separate, overlap-free estimates of the hemodynamic responses associated with the cues and other task events. Instead, the authors modeled the BOLD signal on each trial using one predictor that conflated the cue, the go/change signals, and the response to these signals. This, together with the present results, raises the concern that at least some of the contrasts used by Brown and Braver to identify areas sensitive to error likelihood may have been confounded by hemodynamic activity associated with actual task performance. That is, the activation clusters identified by these contrasts may in part reflect performance-related differences (e.g., in terms of conflict, errors, and/or mental effort) between high and low error-likelihood trials, rather than differences in cue-induced error-likelihood predictions. In contrast, our experiments were designed such that the cue contrasts were not confounded by the effects of these performance-related variables.

This explanation of Brown and Braver’s (2005a) results can be detailed further by considering the specific contrasts that they used. Error-likelihood (EL) areas were identified as those areas that showed significant effects ($p < .05$, uncorrected) for all three of the following contrasts (analyzed using only correct trials): (i) change $>$ go; (ii) high EL/change $>$ low EL/change; (iii) high EL/go $>$ low EL/go. Contrast (i) is not sensitive to error likelihood because there were an equal number of high and low EL trials in the go and change conditions. Furthermore, although contrast (ii) compares high and low EL trials, change trials also elicit more conflict in the high EL than in the low EL condition; that is, because the change-signal delay is longer, it is harder for the participant to reverse the response in time. Our results show significant pMFC activation to such conflict, even at a conservative threshold. As we have noted above, contrast (iii) can be argued to yield a relatively pure measure of error likelihood. However, Brown and Braver’s participants were significantly slower in the high EL/go condition than in the low EL/go condition—a strategic effect that in our experiment was counteracted by the task instruction not to delay the response in
21

anticipation of a possible change signal. This suggests that even contrast (iii) may have been
confounded by differences in task performance.3

In Experiments 1 and 2, we compared the regression coefficients associated with the high
and low error-likelihood cues for each of three specific foci in the pMFC: the two foci identified by
Brown and Braver (2005a), and the pMFC region that was most reliably activated by the two cues
in each of our experiments. Even using a highly sensitive significance threshold (α = .05), none of
the comparisons reached significance, consistent with our general failure to find reliable error-
likelihood effects. Nevertheless, it is striking that all of the numerical differences between the
regression coefficients are in a direction that is consistent with the error-likelihood hypothesis. One
possible reason for this pattern of results is that there is a “true” error-likelihood effect in our data,
but that we have insufficient statistical power to detect it. Note that if this is the case, the size and
consistency of this effect are relatively minor compared to the large and consistent effects of other
variables (e.g., degree of conflict, actual performance accuracy) on pMFC activity. An alternative
possibility is that the observed pattern of results reflects collinearity between the cue and
target/response regressors. Although we increased and jittered the cue-target interval to decrease
the collinearity between the cue and target/response regressors to a level that allowed us to reliably
distinguish the associated BOLD responses (see previous paragraph), there was a small degree of
residual collinearity between some of the regressors. As a result, the contrast between the high and
low error-likelihood cues was slightly confounded by differences in target processing on high
versus low error-likelihood trials (e.g., high error-likelihood cues are associated with increased
conflict and more errors). Although this confound is too small to influence statistical outcomes, its
size and direction are sufficient to account for the observed numerical differences between the
regression coefficients. Thus, there may be a common explanation for Brown and Braver’s
significant error-likelihood effects and the nonsignificant but consistent trends in our data: the fact
that measures of error likelihood tend to be confounded with the effects of task-processing variables that are known to modulate pMFC activity.

A possible alternative explanation for our failure to replicate Brown and Braver’s (2005a) results is that learning of the association between cues and errors may have been hampered by the relatively long and variable cue-target interval used in our study (average ~3 sec, compared to 1 sec in Brown and Braver). Computational analyses suggest that for such associative learning to occur, it is necessary that a representation of the cue is maintained in working memory during the delay between the cue and the target (Brown and Braver 2005b). Yet, it seems unlikely that the present results are attributable to inefficient working memory representations of the cue; despite the relatively long cue-target interval, the cue remained on the screen until the onset of the target, thus placing minimal demands on working memory. Note further that the go and change signals contained the same predictive feature as the cue (i.e., blue or white color). Nevertheless, we conducted a final experiment to address the differences between studies with regard to the timing of stimulus events.

**Experiment 3**

In Experiments 1 and 2 we failed to replicate Brown and Braver’s (2005a) finding that pMFC activity is modulated by the error likelihood predicted by arbitrary visual cues. A notable difference between our experiments and the Brown and Braver study concerned the timing of stimulus events. In particular, the cue-target interval and the intertrial interval were substantially longer in our experiments. Although, as discussed above, these procedural changes can easily be justified on methodological grounds, it is possible that they are responsible for the discrepant results. To address this possibility, we replicated Brown and Braver’s experiment, but this time closely following their timing of stimulus events (e.g., a fixed 1-sec cue-target interval). Importantly, in order to obtain overlap-free estimates of the neural responses associated with the
cues and other task events, we utilized the fine temporal resolution of electroencephalography (EEG) to avoid the methodological limitations of fMRI.

Although Brown and Braver (2005a) are not explicit about the predictions of their hypothesis with regard to ERP components, or concerning the exact timing of cue-driven error-likelihood effects, presumably such effects should be reflected in ERP modulations during the cue-target interval. More specifically, although the data allowed us to look for such modulations at any electrode at any time during this interval, we had a clear hypothesis about where to look first. This hypothesis is based on the fact that previous research has established distinct ERP correlates of the pMFC response associated with response conflict (the N2; Van Veen and Carter 2002; Nieuwenhuis and others 2003; Yeung and others 2004; Bekker and others 2005), errors, and negative feedback (the error-related negativity; Holroyd and others 2004a). The N2 is a negative modulation with a peak latency around 250 ms, and is typically largest at frontocentral electrodes, as one would expect of a pMFC-driven ERP modulation. If we assume that phasic pMFC responses elicited by error-likelihood cues have a similar timing as those elicited by events typically associated with an N2, then we can expect differential pMFC responses to high and low error-likelihood cues to be apparent as voltage differences in the N2 time range following the cue. Therefore, we specifically focused our analyses on possible modulations during this time window. However, our assumption about the timing of error-likelihood effects might be incorrect, so therefore we also scrutinized the frontocentral ERP waveforms for modulations with other temporal properties than the N2. In addition, like in Experiment 2, we also looked for error-likelihood effects during the target/response period by comparing go trials following high vs. low error-likelihood cues.

Two other aspects of Experiment 3 are worth noting. First, we slightly modified the task parameters to increase the difference in error rate between the high and low error-likelihood conditions. In the first two experiments, this difference was smaller than in Brown and Braver’s (2005a) study. This may have resulted in reduced power to detect potential differences in pMFC
activity associated with the two cues. Second, the fast rate of stimulus presentation enabled by the use of EEG allowed us to run a considerably larger number of trials per participant. This, in turn, allowed us to examine learning effects in more detail than in the previous two experiments. In particular, we were interested in the possibility that a possible N2 modulation associated with error likelihood would gradually emerge over the course of the experiment.

Materials and Methods

Participants. Participants were 8 young adults (7 female; average age 23.1 years). All participants were right-handed and all had normal or corrected-to-normal visual acuity. They were paid €15 for a 1.5-h session.

Task and Stimuli. The task and stimuli were the same as in Experiment 2, with the following exceptions. The experimental phase consisted of 864 trials, divided into eight equal blocks, with short breaks in between. Stimuli were presented in color against a black visual display on a monitor placed at eye level at a distance 80 cm from the participant. The timing of the stimulus events replicated that of Brown and Braver (2005a). The error-likelihood cue was presented for a fixed duration of 1000 ms. The go and change signals remained visible until a response deadline of 1000 ms after go signal onset, and the intertrial interval was 250 ms. The staircase tracking algorithm aimed at an error rate of 4% in the low error-likelihood condition, by incrementing the CSD by 2 ms for each correct low-error/change trial, and decrementing the CSD by 50 ms for each incorrect low-error/change trial. The CSDs for both error-likelihood conditions were initialized at 250 ms. Unlike in Experiment 2, but following Brown and Braver (2005a), participants were not given any information about the error-likelihood cues, and were not explicitly instructed to attend to the cues. An informal exit interview revealed that during the experiment 2 out of 8 participants became aware of the relationship between cue identity and task difficulty.
**EEG data acquisition and analysis.** EEG recordings were taken from 15 Ag/AgCl electrodes embedded in a fabric cap (Electro-Cap International, Inc., Eaton, OH), referenced to the left mastoid: F3, Fz, F4, FC3, FCz, FC4, C3, Cz, C4, CP3, CPz, CP4, P3, Pz, and P4. During offline analysis, all signals were re-referenced to the algebraic mean of both mastoids. The electro-oculogram (EOG) was recorded from electrodes placed above and below the left eye, and from electrodes placed on the outer canthi of each eye. All electrode impedances were kept below 10 kΩ. The EEG signals were amplified (Synamps, band-pass filter 0.1-70 Hz), and digitized at 250 Hz.

Single-trial epochs were extracted offline for a period from 200 ms before until 1000 ms after the critical event. Standard Neuroscan (Neurosoft Inc., Sterling VA, USA) analysis procedures were used to correct for EOG artifacts and to discard trials with recording artifacts. Then, for each participant and each condition of interest, the EEG epochs were averaged with respect to cue onset, go-signal onset, and change-signal onset. A baseline, computed as the average signal activity across the 200 ms prior to the stimulus, was subtracted for each ERP. For each participant, the amplitude of ERP components/segments of interest was defined as the average signal value in a carefully chosen time window following stimulus onset (see Results). The analyses focused on electrode Cz, an electrode often used for measuring the N2 and other ERP reflections of pMFC activity (e.g., Kok 1986).

**Results**

**Behavior.** The average error rates on high and low error-likelihood change trials were 51.3% and 14.1%, respectively, $F(1,9) = 211.4, p < .001$. The corresponding average correct RTs on change trials were 382 ms and 387 ms, $F < 1$. Errors on go trials were rare (~1%). The RTs on correct go trials associated with high (379 ms) and low error-likelihood cues (377 ms) were similar, $F < 1$. The mean CSDs on high and low error-likelihood trials were 171 ms and 74 ms, respectively.
As in Experiment 2, we examined whether participants adjusted their behavior in response to various trial types. The sequential-effect analyses indicated that RTs on correct go trials were slower if the immediate preceding trial was a successful change trial (430 ms) as compared to a correct go trial (357 ms), $F(1,7) = 38.7, p < .0001$. Go RT was reliably modulated by the error likelihood of the previous successful change trial (high 440 ms, low 424 ms), $F(1,7) = 16.8, p = .005$, but not by the error likelihood of the previous correct go trial (high 358 ms, low 356 ms), $F < 1$.

**ERPs.** Figure 2A shows the grand-average ERP waveforms elicited by high and low error-likelihood cues. The two waveforms are essentially overlapping. The small amplitude difference in the 600-1000 ms interval was not reliable, $F < 1$. Most importantly, although a hint of a negative component is visible in the 250-325 post-cue interval, if anything the signal is slightly more negative following the low than following the high error-likelihood cues. To examine possible effects of learning we quantified the average signal amplitudes in the aforementioned time window and plotted these as a function of time-on-task (Figure 2B). As is evident in this figure, there was little difference between the two cues across the experiment. A repeated-measures ANOVA with error likelihood (high vs. low) and time-on-task (blocks 1-2, 3-4, 5-6, and 7-8) as within-subject factors yielded no significant main or interaction effects, all $F$s < 1. Essentially the same pattern of results was obtained for other frontocentral electrodes.

As in Experiment 2, we looked for possible error-likelihood effects during later stages of the trial. Figure 2C presents the ERP waveforms elicited by go signals (for go trials only) as a function of whether the preceding cue indicated a high or low error likelihood. The two waveforms are almost identical, suggesting that pMFC activity during the response period was not modulated by error likelihood.

To investigate whether we obtained the typical ERP pattern associated with the necessity to stop and reverse an ongoing response, we conducted an additional analysis in which we examined
the ERP waveforms elicited on go trials, successful change trials, and unsuccessful change trials, separately for the high (Figure 2D) and low (Figure 2E) error-likelihood condition. In line with previous research (e.g., Schmajuk and others 2005; Ramautar and others 2006), both successful and unsuccessful change trials were associated with a negativity in the N2 time range that was not present on go trials. Both negativities were characterized by a broad midline scalp distribution. To quantify and compare these N2 components, we computed the average signal values in the window 175-250 ms following change-signal onset, and compared these between conditions using paired t-tests (one-tailed). In the high error-likelihood condition, successful change trials ($M = 4.7 \mu V, SD = 2.7 \mu V$) were associated with a larger N2 than go trials ($M = 6.6 \mu V, SD = 3.5 \mu V$), $t(7) = 2.1, p = .04$, although admittedly the modulation starts well before the N2 time window. Furthermore, unsuccessful change trials ($0.9 \mu V, SD = 3.6 \mu V$) were characterized by a larger N2 than successful change trials, $t(7) = 3.8, p < .005$. In the low error-likelihood condition, successful ($M = 5.5 \mu V, SD = 2.2 \mu V, t(7) = 1.7, p = 0.06$) and unsuccessful ($M = 5.3 \mu V, SD = 3.0 \mu V, t(7) = 1.6, p = 0.08$) change trials were associated with a larger N2 than go trials ($M = 7.0 \mu V, SD = 2.4 \mu V$), although both effects just missed significance. Thus, especially when taken together, the results from the high and low error-likelihood condition corroborate previous studies in demonstrating N2 modulations on successful and unsuccessful change trials (Van Boxtel and others 2001).

Discussion

The results from Experiment 3 indicate that error-likelihood differences were not associated with an N2 modulation in the cue-locked ERP waveforms. Indeed, the ERP waveforms elicited by the two cues were essentially overlapping, indicating that we also found no evidence for error-likelihood modulations outside the N2 time window. Similar results have been reported by Holroyd and Coles (2002), who found that stimuli that were always associated with negative feedback (i.e., regardless of the response to those stimuli) did not elicit an N2. Therefore, on the assumption that
pMFC responses to error-likelihood cues are measurable in the ERP, the current findings would seem inconsistent with the prediction of the error-likelihood hypothesis that the cues elicit pMFC responses that are proportional to the error likelihood associated with each cue. The results from Experiment 3 also seem inconsistent with the possibility that an error-likelihood effect on pMFC activity occurred in conjunction with the go signal and associated response.

In contrast, and in line with the fMRI results of Experiment 2, the ERP waveforms showed suggestive evidence of increased pMFC activity associated with errors (unsuccessful vs. successful change trials; in the high error-likelihood condition only). That is, the corresponding contrast revealed clear signs of an N2 modulation, replicating prior research using the stop task (Van Boxtel and others 2001; Ramautar and others 2006). The observed N2 modulation on successful change vs. go trials has also been suggested to reflect pMFC activity, elicited by the increased conflict on change trials. However, it is important to note that contrasts between change and go trials are typically confounded by the presence of an additional stimulus (i.e., the change signal) on change trials. Therefore, it is hard to exclude the possibility that the increased N2 on change trials reflects an evoked response to the change signal, rather than the presence of increased response conflict. To our knowledge, there is only one study with the stop-change paradigm that has controlled for this confound, by including a control condition in which the stop signal had to be ignored. This study found substantially increased N2 amplitudes in the stop vs the control condition (Schmajuk et al., 2005), suggesting that the N2 did not reflect an evoked response to the stop signal. Nevertheless, in the present experiment we did not control for this confound, because we wanted to closely replicate the task design used by Brown and Braver (2005a). Consequently, we need to be cautious in interpreting the increased N2 on successful change trials. However, at the very least, this finding indicates that the experiment had sufficient power to detect the ERP modulations typically obtained in the stop-change paradigm, which places in perspective our failure to detect ERP correlates of
error likelihood. Thus, when taken together, the results of Experiment 3 seem to pose a challenge for the error-likelihood hypothesis.

**General Discussion**

In the present research we tested the error-likelihood hypothesis (Brown and Braver 2005a), an elegant new hypothesis that attempts to resolve debate about the function of the pMFC by suggesting how previous theories can be integrated in one overarching account. According to the error-likelihood hypothesis, the pMFC learns to predict the likelihood of an error occurring in a given task context, and uses this information to alert other brain systems that cognitive control needs to be increased. We conducted three experiments, each with one condition in which participants made many errors, and another condition in which they made relatively few errors. Furthermore, each trial started with the presentation of an arbitrary visual cue that signaled whether the likelihood of an error on that particular trial was high or low. In each experiment, the critical prediction of the error-likelihood hypothesis was that pMFC activity should be systematically influenced by the error-likelihood signaled by the cue.

The main results are clear-cut: None of the experiments provided support for the prediction of the error-likelihood hypothesis. More specifically, we found no evidence for an effect of cue-signaled error likelihood on pMFC activity, nor any indication that such an effect developed with increasing task experience. This was even the case when the statistical threshold was lowered to quite liberal levels. This does not imply that the error-likelihood cues did not activate the pMFC. Indeed, in Experiments 1 and 2 we found that both cues elicited significant responses in the caudal extent of the pMFC (possibly reflecting non-specific arousal; Downar and others 2000), but they did so to a similar extent. We also found no evidence for robust error-likelihood effects in other brain areas, or during the response period of the trial. These results constitute a failure to replicate Brown and Braver’s (2005a) fMRI results and are problematic for the error-likelihood hypothesis.
Our negative findings concerning the effect of error likelihood are particularly striking when regarded in the context of the reliable increases in pMFC activity that we observed in relation to task difficulty (Experiment 1), negative feedback (Experiment 1), response conflict (Experiment 2), and errors (Experiments 2 and 3). These effects are consistent with numerous published studies (Botvinick and others 2004; Holroyd and others 2004; Ridderinkhof and others 2004), and have been an important source of evidence for previous theories of pMFC function. Thus, our results suggest that although the pMFC is sensitive to response conflict, errors, and other undesired states, it is insensitive to arbitrary cues that predict these states.

Why did we not manage to replicate Brown and Braver (2005a)? One possible explanation is that the participant groups partaking in the two studies differed with regard to some characteristic that affects pMFC function. For example, task-related responses in pMFC are known to be influenced by differences in various personality dimensions (e.g., Gray and others 2005), and the two groups might differ along one of these dimensions. Unfortunately it is hard to assess this possibility because no measures of personality or similar characteristics were obtained in either study.

Another tentative explanation involves the number of trials it takes for the hypothesized error-likelihood effect to become manifest. Obviously, associative learning of the relation between the cues and corresponding error rates requires experience with the task. As a consequence of the longer trial durations, we ran considerably fewer trials in Experiments 1 (180) and 2 (240) than in Brown and Braver’s experiment (~420 trials analyzed), and perhaps the error-likelihood effect requires more trials to develop. However, there are at least three arguments against this interpretation. First, to counteract the effects of the reduced trial numbers (and unlike Brown and Braver), we explicitly informed the participants in Experiments 1 and 2 about the significance of the cue—that is, its relation with the type of search display in Experiment 1, and with the change-signal delay in Experiment 2. We assumed that by informing them in advance, participants would
require less task experience to acquire the mapping between cues and error rates. This assumption receives some support from our informal exit interviews, which suggested that two thirds of the participants in Experiment 1 became aware of the relationship between the cues and task difficulty/error rates (compared to half of the participants in Brown and Braver). Second, in Experiment 3 we ran 864 trials and—even in the last quarter of the experiment—found no indication of an error-likelihood effect in the ERP waveforms. We also found no evidence of such an effect in the second half of Experiments 1 and 2. Finally, Brown and Braver’s own results are somewhat ambiguous with regard to the effect of task experience. In one region of interest (in the pre-SMA), the error-likelihood effect did not develop until the second half of the experiment (i.e., block 3 of 4), suggesting that the forming of error-likelihood predictions requires a large number of trials. However, in the second region of interest (in the ACC), the error-likelihood effect did not significantly increase from block 1 to 4 (J.W. Brown, personal communication, August 5, 2005), which implies that sensitivity to the cues emerged early in the experiment. In sum, although it is premature to discount an explanation of our results in terms of the amount of task experience, such an explanation seems at odds with various aspects of our results and those of Brown and Braver.

Alternatively, our results might represent an indication that the error-likelihood hypothesis is incorrect. As we have argued, the empirical evidence put forward by Brown and Braver (2005a) as support for their hypothesis is somewhat undermined by methodological concerns: Their task design precluded the effective separation of the pMFC responses to the cues and to other aspects of the task, suggesting that at least some of the critical contrasts may have measured pMFC modulations associated with variables other than cue-signaled error likelihood. Considering this possibility along with the present empirical results, the current status of the error-likelihood hypothesis would appear to be that it is an intriguing idea in need of additional supporting evidence.
Acknowledgements

This research was supported by the Netherlands Organization for Scientific Research (S.N.), and the National Institutes of Mental Health (NIMH) grants MH062196 (M.M.B.) and MH069047 (G.H.). We thank Niels Alting von Geusau and Dirk Heslenfeld for technical assistance, and Josh Brown and Todd Braver for helpful comments on an earlier version of this manuscript.

Footnotes

1 Although Brown and Braver (2005a) focused their error-likelihood hypothesis on the ACC, one of the two mediofrontal areas that showed effects consistent with the hypothesis could arguably be described as falling within the pre-SMA, rather than in the ACC, as Brown and Braver suggest.

2 It should be acknowledged that negative feedback occurred less frequently than positive feedback and that hence the effects of feedback valence were potentially confounded with effects of stimulus frequency. Based on empirical and theoretical work (Holroyd and others 2003), it is likely that the observed feedback effect on pMFC activity reflects in part the interaction of valence and frequency, rather than just the main effect of one of these factors.

3 As evidence against this possibility, Brown and Braver (2005a) showed that this contrast yielded the same results when RT was included as a nuisance covariate in the general linear model. Even so, the only stringent criterion for identification as an error-likelihood area was contrast (iii) with a significance threshold $\alpha = .05$ (uncorrected). Given the use of this liberal threshold, it is possible that the obtained error-likelihood effects reflect false positives.
However, note that in Experiment 2 and in Brown and Braver’s (2005a) experiment, the reported error rates are based on change trials, which comprised only one third of the trials. In fact, the error-likelihoods associated with the two cues were substantially lower, because the cues were also presented on go trials, on which participants made almost no errors. The error rates reported for Experiment 1 are based on all trials.

Yet another pattern of results was obtained with computer simulations of a model implementing the error-likelihood hypothesis (Brown and Braver 2005a). These simulations suggest that the error-likelihood effect should emerge after very little experience, and then steadily increase with learning.
References


Table 1. fMRI results in Experiment 1

<table>
<thead>
<tr>
<th>Area</th>
<th>Left/Right</th>
<th>Volume (mm$^3$)</th>
<th>X</th>
<th>Y</th>
<th>Z</th>
<th>Max t value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>High &gt; Low error-likelihood cue</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Conjunction search &gt; Popout search$^a$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ACC / pre-SMA</td>
<td>Right</td>
<td>18221</td>
<td>3</td>
<td>19</td>
<td>37</td>
<td>13.96</td>
</tr>
<tr>
<td>Insula</td>
<td>Right</td>
<td>7269</td>
<td>39</td>
<td>18</td>
<td>8</td>
<td>12.18</td>
</tr>
<tr>
<td>Insula</td>
<td>Left</td>
<td>6386</td>
<td>-34</td>
<td>17</td>
<td>5</td>
<td>14.28</td>
</tr>
<tr>
<td>Intraparietal sulcus</td>
<td>Left</td>
<td>1789</td>
<td>-30</td>
<td>-51</td>
<td>38</td>
<td>9.60</td>
</tr>
<tr>
<td>Popout search &gt; Conjunction search$^a$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Posterior cingulate</td>
<td>Left</td>
<td>3092</td>
<td>-1</td>
<td>-54</td>
<td>25</td>
<td>8.24</td>
</tr>
<tr>
<td>Rostral anterior cingulate (subgenual)</td>
<td>Left</td>
<td>1446</td>
<td>-2</td>
<td>40</td>
<td>-4</td>
<td>8.93</td>
</tr>
<tr>
<td><strong>Negative &gt; Positive feedback</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Superior frontal gyrus</td>
<td>Right</td>
<td>131</td>
<td>6</td>
<td>15</td>
<td>63</td>
<td>6.67</td>
</tr>
<tr>
<td>ACC$^*$</td>
<td>Right</td>
<td>1041</td>
<td>5</td>
<td>29</td>
<td>28</td>
<td>5.35</td>
</tr>
<tr>
<td>Pre-SMA$^*$</td>
<td>Right</td>
<td>1528</td>
<td>9</td>
<td>8</td>
<td>65</td>
<td>6.67</td>
</tr>
<tr>
<td>Pre-SMA$^+$</td>
<td>Right</td>
<td>162</td>
<td>4</td>
<td>7</td>
<td>52</td>
<td>4.93</td>
</tr>
<tr>
<td><strong>Positive &gt; Negative feedback</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Middle frontal gyrus</td>
<td>Left</td>
<td>300</td>
<td>-25</td>
<td>17</td>
<td>55</td>
<td>6.72</td>
</tr>
<tr>
<td>Globus pallidus / putamen</td>
<td>Right</td>
<td>151</td>
<td>13</td>
<td>5</td>
<td>0</td>
<td>6.21</td>
</tr>
</tbody>
</table>

Note: All regions are $p < 0.0005$ (uncorrected, voxel contiguity = 120 mm$^3$), except for those indicated with an ‘*’, which were identified by a region-of-interest (ROI) analysis with $p < 0.005$ (uncorrected). Areas in the posterior medial frontal cortex (pMFC) are boldfaced. Pre-SMA = pre-supplementary motor area. ACC = anterior cingulate cortex.

$^a$ For these contrasts the voxel contiguity threshold was increased to 1000 mm$^3$ to limit the number of identified areas.
Table 2. fMRI results in Experiment 2

<table>
<thead>
<tr>
<th>Area</th>
<th>Left/Right</th>
<th>Volume (mm$^3$)</th>
<th>X</th>
<th>Y</th>
<th>Z</th>
<th>Max t value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>High &gt; Low error-likelihood cue</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>High go &gt; Low go</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Successful change &gt; Go$^a$</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre-SMA</td>
<td>Right</td>
<td>158</td>
<td>6</td>
<td>17</td>
<td>55</td>
<td>6.59</td>
</tr>
<tr>
<td>Inferior frontal gyrus</td>
<td>Right</td>
<td>326</td>
<td>50</td>
<td>14</td>
<td>6</td>
<td>7.02</td>
</tr>
<tr>
<td>Inferior frontal gyrus</td>
<td>Left</td>
<td>162</td>
<td>-49</td>
<td>13</td>
<td>7</td>
<td>6.73</td>
</tr>
<tr>
<td>Insula</td>
<td>Right</td>
<td>178</td>
<td>32</td>
<td>20</td>
<td>4</td>
<td>6.01</td>
</tr>
<tr>
<td>Insula</td>
<td>Left</td>
<td>669</td>
<td>-34</td>
<td>19</td>
<td>5</td>
<td>6.97</td>
</tr>
<tr>
<td>Precentral gyrus</td>
<td>Right</td>
<td>277</td>
<td>38</td>
<td>1</td>
<td>35</td>
<td>8.23</td>
</tr>
<tr>
<td><strong>Unsuccessful &gt; Successful change</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ACC</td>
<td>Right</td>
<td>309</td>
<td>3</td>
<td>38</td>
<td>25</td>
<td>5.87</td>
</tr>
<tr>
<td>Medial frontal gyrus (BA8)</td>
<td>---</td>
<td>169</td>
<td>0</td>
<td>22</td>
<td>43</td>
<td>5.83</td>
</tr>
<tr>
<td>Insula</td>
<td>Right</td>
<td>186</td>
<td>36</td>
<td>17</td>
<td>16</td>
<td>6.97</td>
</tr>
<tr>
<td>Inferior parietal lobule</td>
<td>Left</td>
<td>272</td>
<td>-51</td>
<td>-35</td>
<td>25</td>
<td>7.06</td>
</tr>
</tbody>
</table>

Note: All regions are $p < 0.0005$ (uncorrected, voxel contiguity = 120 mm$^3$). Areas in the posterior medial frontal cortex (pMFC) are boldfaced. Pre-SMA = pre-supplementary motor area. ACC = anterior cingulate cortex. BA = Brodmann area.

$^a$For this contrast, brain areas with $Y < -30$ are not reported; activations in posterior areas likely reflect the visual processing of the change signal, which was present on successful change trials but not on go trials.
Figure Captions

Figure 1. Example sequence of stimulus events in Experiment 1. See text for details about actual colors. ‘Fout’ is Dutch for ‘error’. ITI = intertrial interval.

Figure 2. Results from Experiment 3. All graphs show data from electrode Cz. (A) Grand-average cue-locked ERP waveforms associated with high and low error-likelihood cues. The two dotted lines mark the time window to which summary measures and statistical analyses were constrained. (B) Average signal amplitude as a function of error-likelihood condition and time-on-task. (C) Grand-average ERP waveforms elicited by go signals preceded by high and low error-likelihood cues (go trials only). (D/E) Grand-average ERP waveforms associated with various trial types (unsuccessful change, successful change, go), presented separately for high and low error-likelihood trials. To construct the go-trial ERP, the EEG signals were aligned to the onset of a “virtual change signal”: the moment a change signal would have occurred if the trial were a change trial. The waveforms in (D) and (E) were low-pass filtered (<20 Hz) for presentation purposes only.
Figure 1

- Error-likelihood cue: 1.0 – 6.0 sec
- Search display: X msec (variable)
- Mask display: (2000 – X) msec
- Feedback: 500 msec
- ITI: 1.5 – 15.0 sec