The conflict adaptation effect: It's not just priming

MARKUS ULLSPERGER
Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany

and

LAUREN M. BYLSMA and MATTHEW M. BOTVINICK
University of Pennsylvania, Philadelphia, Pennsylvania

Analyses of trial sequences in flanker tasks have revealed cognitive adaptation, reflected in a reduced interference effect following incompatible trials (Gratton, Coles, & Donchin, 1992). These effects have been explained on the basis of the response conflict monitoring model of Botvinick, Braver, Barch, Carter, and Cohen (2001), who proposed that preceding response conflict triggers stronger top-down control, leading to performance improvements on subsequent trials of similar context. A recent study (Mayr, Awh, & Laurey, 2003) has challenged this account, suggesting that the behavioral adaptations are confined to trial sequences of exact trial repetitions and can therefore be explained by repetition priming. Here, we present two experiments in which the sequential dependency effect was present even on trial sequences that did not involve stimulus repeats. We discuss the data with respect to the conflict-monitoring and repetition-priming accounts.

In a recent work, Botvinick, Braver, Barch, Carter, and Cohen (2001) proposed that cognitive control is modulated, in part, on the basis of a process referred to as *response conflict monitoring*. According to this account, increased top-down control over information processing is triggered by the occurrence of response competition. In addition to certain neuroscientific data, the conflict monitoring theory is based on a set of behavioral phenomena that appear to reflect online reactive adjustments in control. A prominent example is provided by Gratton, Coles, and Donchin (1992), who reported evidence of a sequential dependency effect in the Eriksen flanker task (Eriksen & Eriksen, 1974). The flanker task calls for a left or right response based on the identity of a centrally presented target symbol. This target is surrounded by distractor flanker symbols, which themselves map to responses that are either compatible or incompatible with the required response. Gratton et al. investigated the effects of trial-type transitions (compatible–compatible [C–C], compatible–incompatible [C–I], incompatible–compatible [I–C], and incompatible–incompatible [I–I]) and showed that the interference effect (reaction time [RT] for incompatible trials minus RT for compatible trials) was reduced following incompatible trials. The occurrence of an incompatible trial thus appeared to enhance target processing and/or suppress flanker processing on the following trial. As Botvinick and colleagues ar-

gued (Botvinick et al., 2001; Botvinick, Nystrom, Fissell, Carter, & Cohen, 1999), this finding appears to provide an example of the reactive adjustments in control posited by the conflict monitoring hypothesis: Incompatible trials involve response conflict, and it is this that causes them to be associated with a subsequent intensification of top-down control.

A challenge to this account was recently put forth by Mayr, Awh, and Laurey (2003). They suggested that the effect reported by Gratton et al. (1992) might simply reflect repetition priming. Note that one way of describing that effect is as a shortening of RTs in trials in which the stimulus type (compatible vs. incompatible) is the same as it was on the preceding trial. Note further that in the usual version of the task, such trial type repeats also frequently involve a repetition of the entire stimulus (e.g., $> > < > > \rightarrow > > < > >$). Taking both of these points into account, Mayr et al. suggested that stimulus repetition itself might be responsible for the faster RTs seen with trial type repeats, simply as a consequence of repetition priming.

Mayr et al. (2003) presented evidence for this account from two experiments. In the first, participants performed a version of the flanker task using left- and right-facing arrow heads. Although their performance displayed the effect originally described by Gratton et al. (1992), this effect was limited to trials in which the target item was the same as it was on the preceding trial. Trials in which there was no repeat showed no such effect. A second experiment, in which stimulus elements never repeated from one trial to the next, also failed to show the effect described by Gratton et al.

The findings reported by Mayr et al. (2003) are surprising for a number of reasons. First, they contrast with the results of Gratton et al. (1992), who addressed the

priming explanation for their findings by showing that there was no significant difference between target alternation and target repetition trials.[1] Moreover, in their Experiment 3 Gratton et al. showed that similar effects can be obtained by modulating expectancy for incompatible trials using an arbitrary precue, thus eliminating the relationship with previous stimuli and responses as a factor. Second, there is evidence that conflict adjustment effects can be observed, independent of priming, in other tasks (e.g., the Stroop task [Kerns et al., 2004] and the Simon task [Sohn & Carter, 2003]; see the Discussion section). Given these considerations, it seems important to determine the extent to which the findings of Mayr et al. are replicable and, in particular, the extent to which they generalize across task implementations.

We report here the results of two experiments that speak to this issue. In each, a conflict adaptation effect was observed in the flanker task, in a form not attributable to priming. The first experiment involved a reanalysis of previously collected data based on the conventional arrowhead flanker task. The second was a new experiment, in which an attempt was made to minimize repetitions of stimulus elements from one trial to the next by using a larger stimulus set.

EXPERIMENT 1

Method

Participants. Nineteen University of Leipzig undergraduates (9 female) took part in the study. Their ages ranged from 19 to 31 years. The participants provided informed consent prior to the beginning of the experiment and received an hourly base rate for their participation.

Procedure. The stimuli, which were presented using the ERTS software package (BeriSoft, Frankfurt), consisted of a central left- or right-facing arrow with two flanker arrows above and two below. The arrows were 0.46º tall and 1.08º wide, and the flankers were presented 0.52º and 1.04º above and below the screen center. Targets were presented for 30 msec together with the flankers, which first appeared 80 msec earlier than the targets. The left and right index fingers were used for buttonpress responses according to the direction indicated by the target arrow. Responses had to be given before a response deadline (377 msec). When the response was given after the deadline had elapsed, the participants received feedback indicating that they should speed up their responses. The participants performed two 264-trial blocks containing randomly sequenced 50% incompatible and 50% compatible trials. The intertrial interval (ITI) varied randomly between 5,000 and 6,000 msec in steps of 500 msec. The blocks differed with respect to instructions, which were reinforced by financial incentives. In both blocks, correct responses were rewarded (€0.07). In one block, response accuracy was instructed and errors were penalized more (−€0.40) than late responses (−€0.20). In the other block, response speed was stressed in the instructions and, in consistency with this, late responses were penalized more (−€0.40) than errors (−€0.20). Block sequence was counterbalanced across participants. Between blocks, the participants had a short break, after which they received the instructions for the next block. Behavioral data were collected and event-related potentials were recorded during the experiments. The electrophysiological data, irrelevant to the issue discussed here, is reported elsewhere (Ullsperger & Szymanowski, 2004).

Results

Four-way analyses of variance (ANOVAs) were conducted for both error rates and hit RTs, with factors of current trial type (compatible vs. incompatible) and previous trial type (compatible vs. incompatible), a third factor (target type) coding for the relationship of the target item on the current trial to the target on the previous trial (target-repeat vs. target-alternation), and a fourth factor coding for instruction (speed vs. accuracy). Trials following errors were excluded from the analyses. Figure 1 depicts the mean RTs and error rates as a function of response conflict on the current trial and resolved conflict on the preceding trials, broken down by the target-type and the instruction factors. The results of the ANOVAs are shown in Table 1.

Both analyses yielded significant main effects of current trial type (the standard flanker effect) and target type. Moreover, the main effect of instruction for error rate confirms that fewer errors were made when accuracy was emphasized, whereas RTs tended to be longer. The main effect of previous trial type was significant for error rate but not for RT. More to the point, the conflict adaptation effect was evident in an interaction between current and previous trial types for both analyses. Critically, and in contrast with the findings of Mayr et al. (2003), the conflict adaptation effect was observed even on target alternation trials. Follow-up ANOVAs focusing only on these trials yielded a significant interaction between current and previous trial types [for RT, $F(1,18) = 10.89, p < .01$; for error rate, $F(1,18) = 9.92, p < .01$] as well as a main effect of previous trial type for error rate [$F(1,18) = 6.33, p < .05$]. No significant interactions between previous trial type and instruction or between previous trial type and target type were found. In follow-up comparisons, the interaction between current and previous trial types was addressed, revealing that error rates were lower in I–I trial sequences than in C–I trial sequences [main effect of previous trial type: $F(1,18) = 10.05, p < .01$]. For RT, an interaction between previous trial type and instruction was found [$F(1,18) = 8.88, p < .01$], reflecting the fact that RT was shorter in I–I trial sequences than in C–I trial sequences in the accuracy condition ($p < .0001$) but not in the speed condition ($p = .41$). In compatible trials, RTs showed a tendency to be shorter in C–C trial sequences than in I–C trial sequences [$F(1,18) = 3.58, p = .07$]. Due to floor effects, for error rates in compatible trials no significant sequence effect was found ($p = .74$).

EXPERIMENT 2

Method

Participants. The participants included 8 University of Pennsylvania undergraduates (4 female) who had responded to an electronic newsgroup posting. Their ages ranged from 18 to 21 years. The participants provided informed consent prior to the beginning of the experiment and received an hourly base rate for their participation.
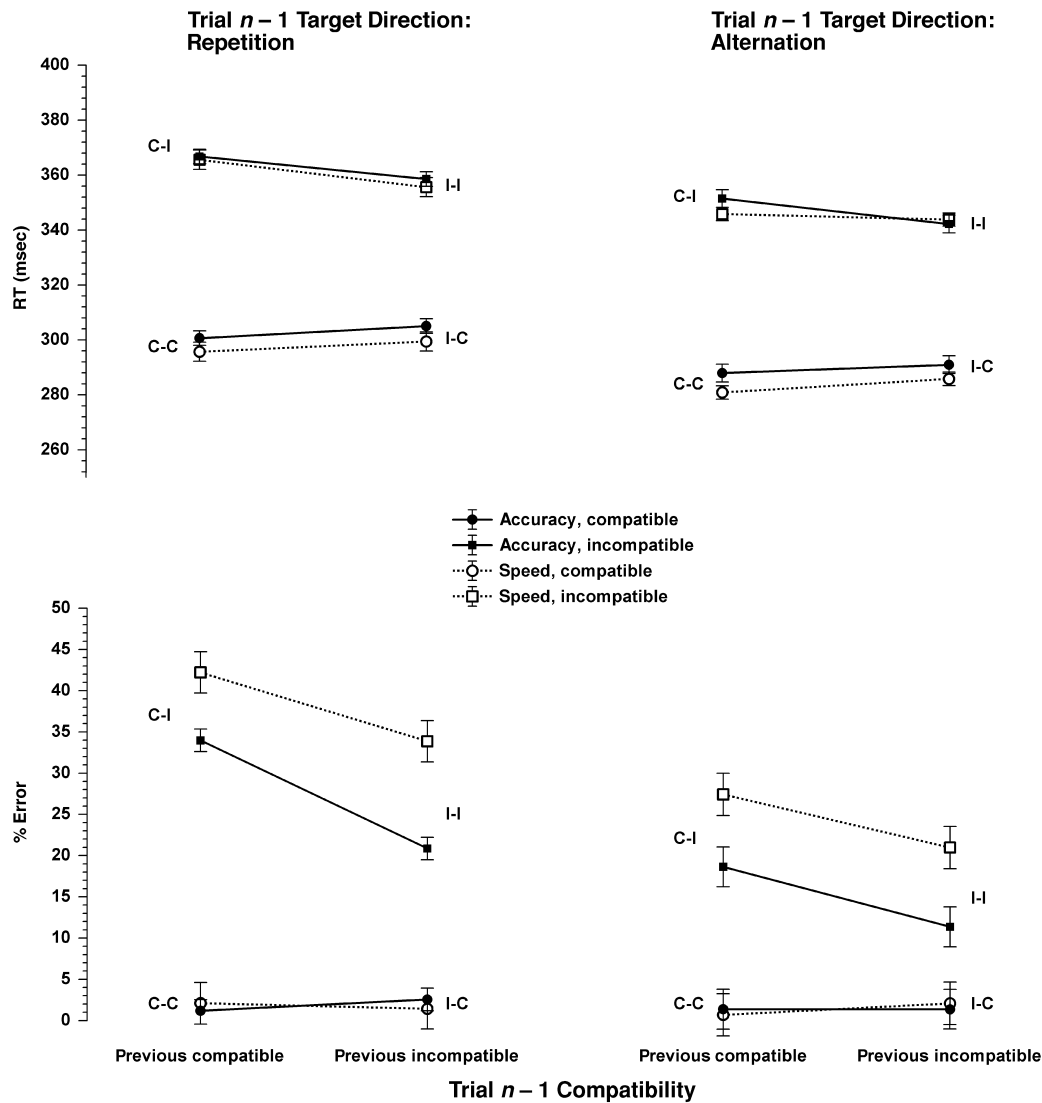
Figure 1. Mean reaction times (RTs, upper panel) and error rates (lower panel) as a function of response conflict on trial *n* and resolved conflict on the preceding trial *n* − 1, broken down by the target transition factor (left, repetition; right, alternation) and by instruction (solid lines, accuracy instruction; dotted lines, speed instruction). Error bars show standard errors of the means, computed to partial out between-sessions variance (following Loftus & Masson, 1994).

Procedure. The experiment was conducted on a standard desktop computer using the E-Prime software package (Psychological Research Tools, Pittsburgh, PA). Each stimulus consisted of a six-digit array presented at the center of the computer monitor. Each array consisted of a central target digit (between 1 and 9, inclusive) and four identical flanker digits (also between 1 and 9, inclusive), two on either side of the target. The target item was underscored. Stimuli could be either compatible (flanker digits identical to target digit) or incompatible (flanker digits different from target). Beyond this constraint, component digits were selected randomly. In response to each stimulus, the participant was instructed to indicate the target digit by pressing the appropriate key on the computer's numeric keypad. Trials began with a 50-msec warning tone followed by a 1,000-msec preparatory period. Next, the target appeared and remained on the screen for 100 msec. The ITI, measured from the participant's response, varied randomly between 3,500 and

5,500 msec (in steps of 250 msec). Within each block, compatible and incompatible trials were intermixed randomly and in equal proportions. Each participant participated in one practice and three experimental sessions, each composed of 15 blocks of 80 trials. Each session lasted approximately 2 h, and the participants were given the opportunity to take a 10-min break at the halfway point. The participants participated in no more than one session per day and completed all sessions within 1 week.

Results

Data analysis focused on trials on which none of the stimulus elements had appeared on the preceding trial. Trials following errors were discarded, as were the first five trials in each block. Average RTs (for correct responses) and error rates are shown in Figure 2.

Table 1
Results of the Repeated Measures ANOVAs for
Error Rates and Hit Reaction Times (RTs)

Factors and Interactions	df	Error Rate F	Hit RT F
Current trial type	1,18	106.82*	333.46*
Previous trial type	1,18	15.22*	1.61
Target type	1,18	15.63*	10.99*
Instruction	1,18	14.45*	3.48†
Current trial type × instruction	1,18	18.85*	1.74
Current trial type × target type	1,18	20.11*	0.60
Current trial type × previous trial type	1,18	21.06*	33.89*

†$p < .1$.   *$p < .01$.

A repeated measures (within-blocks) ANOVA on RT was conducted, with current trial type and previous trial type as factors. A main effect of current trial type was present [$F(1,23) = 162.17$, $p < .001$], confirming the presence of the basic flanker effect. More important, a significant interaction was found between current and previous trial types [$F(1,23) = 6.36$, $p = .019$]. The details of this interaction were consistent with the conflict adaptation effect; the flanker effect (RT for incompatible trials minus RT for compatible trials) was 26% smaller following incompatible trials than following compatible ones (19 vs. 26 msec). The main effect of previous trial type did not reach significance [$F(1,23) = 2.68$, $p = .12$]. A comparable ANOVA on error rate yielded a main effect of current trial type [$F(1,23) = 21.31$, $p < .001$] and a main effect of previous trial type [$F(1,23) = 4.57$, $p < .05$], reflecting slightly lower error rates for both compatible and incompatible trials when these followed incompatible trials.

Two comments are merited concerning these results. The first pertains to the size of the RT effect. It is important to note, in this connection, that the basic interference effect was small—an anticipated consequence of increasing the response set from the usual two to nine. Because the sought adjustment effect would not have been expected to surpass the basic interference effect, the former effect was predicted to be well below 25 msec. It should not be surprising, therefore, that the absolute size of the adjustment effect was small, and this does nothing to undermine its theoretical implications. Indeed, as a proportion of the overall interference effect, the adjustment was similar in magnitude to that reported by Gratton et al. (1992).

A second comment pertains to error rates. The presence of a main effect of previous trial type suggests that two forms of conflict adaptation may have been occurring: (1) an increased focus on the central target item, evident in the interaction effect seen in the RT data, and (2) a speed–accuracy trade-off, reflected in overall lower error rates (and the trend toward slower responses[2]) following incompatible trials. This speed–accuracy trade-off can be seen as an adaptation to the specifics of the task, which, unlike the task used in Experiment 1, required the participants to select among nine response keys. Further analyses suggested that the majority of errors in this experiment were due to inaccurate aiming of manual responses rather than to stimulus-driven interference effects (specifically, 76% of the errors involved striking a key adjacent to the target; in only 3% of incompatible trials did the response match the flanker item). Given the frequency of such spatial inaccuracies,
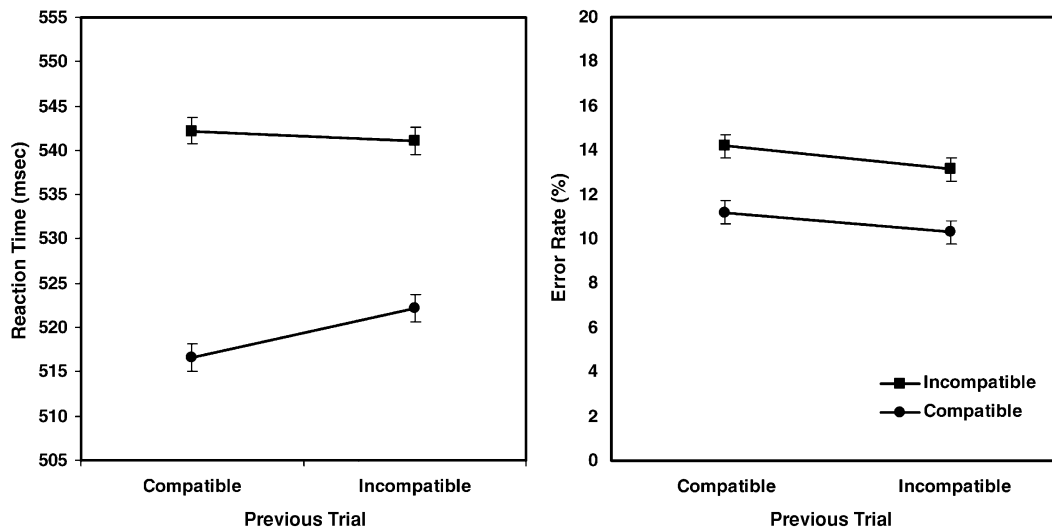

Figure 2. Left: Mean reaction times (RTs) for compatible and incompatible trials, reported separately for trials coming after compatible and incompatible trials. Error bars show standard errors of the means, computed to partial out between-sessions variance (following Loftus & Masson, 1994). Right: Mean error rates for the same trial types.

it may have been useful to trade speed for accuracy following high-conflict trials.

DISCUSSION

In both experiments, significant sequence dependency effects were found—that is, RTs and error rates were lowered on incompatible trials that were preceded by incompatible ones. Because in both cases target repetitions were excluded from analysis, the findings cannot be explained by response priming. Instead, in line with the account put forth by Botvinick et al. (2001), the observed fluctuations in behavior appear consistent with adjustments in control triggered by conflict.

Given all these results, an obvious question is why Mayr et al. (2003) failed to observe a conflict adjustment effect in their two experiments. One possible explanation, which applies to their first experiment, is that the adjustment effect may have been masked by negative priming. Stadler and Hogan (1996) have shown that RTs are unusually high for incompatible flanker stimuli when the locations of target and flanker items are reversed from those of the preceding trial (e.g., $< < > < < \rightarrow >$ $> < > >$). In the analysis presented by Mayr et al., which took only target change trials into account, this negative priming effect would have the result of inflating RTs for I–I trials, possibly masking an underlying conflict adjustment effect. Our own experiments may have avoided this masking effect by minimizing the effect of negative priming. Such priming is unlikely to have played any role in our Experiment 2, in which stimulus elements rarely repeated from one trial to the next. Although such repetitions did occur in Experiment 1, the associated priming effect may have been minimized by the use of relatively long ITIs and brief stimulus presentation times (in contrast with Mayr et al., 2003, in which stimuli remained on the screen until the participant responded, and the response–stimulus interval was 1,000 msec; U. Mayr, personal communication, May 28, 2003). It seems plausible that these parameters minimized negative priming effects, allowing us to avoid the hypothesized masking effect. A further difference with the Mayr et al. study may be important: Gratton et al. (1992) and the present authors used speeded tasks, whereas Mayr et al. applied unspeeded versions, which may have resulted in a decreased need to utilize conflict-based strategy adjustments.

The findings from the second experiment of Mayr et al. (2003) are somewhat harder to explain. In that experiment, trials alternated between stimuli consisting of left- and right-facing arrows and stimuli consisting of upward- and downward-facing arrows. One possibility is that this alternation was treated by participants as a switch between two independent tasks. Such switching has been shown to involve complex effects on cognitive control (Rogers & Monsell, 1995). In view of this, it is not clear that the conflict monitoring model would predict the top-down control modulations across the two

tasks embedded in the Mayr et al. experiment. Recently, Corballis and Gratton (2003) showed that sequence dependence effects in a flanker task do not generalize from stimulus locations in one hemifield to locations in the other hemifield. Similarly, it may be hypothesized that response-conflict-triggered attentional modulations would not generalize from one stimulus orientation to the other.

Our present findings fit well with observations from other tasks. For example, in a study of the Stroop task, Kerns et al. (2004) found less influence of word identity on color naming following incompatible trials than following compatible trials, in consistency with the idea that participants focused more exclusively on the task-relevant color dimension following high-conflict incongruent trials. Sohn and Carter (2003) obtained parallel results in the Simon task—that is, interference effects on trials following incongruent trials were attenuated, again in accordance with the idea that participants focused more on the task-relevant stimulus dimension following high-conflict responses.

The idea that conflict triggers adjustments in control is also supported by data from cognitive neuroscience. Several neuroimaging studies have supported the idea that posterodorsal mesial frontal cortex (pMFC), in the vicinity of the anterior cingulate cortex, responds to the occurrence of response conflicts (e.g., Botvinick et al., 1999; Carter et al., 1998; MacDonald, Cohen, Stenger, & Carter, 2000; Ullsperger & von Cramon, 2001). Recent neuroimaging evidence supports the view that this conflict-related activity is linked to subsequent adjustments in control; Kerns et al. (2004) showed, in the Stroop task, that the strength of trial-specific pMFC activation predicts the degree of Stroop interference (and, by inference, the state of top-down control) on the subsequent trial.

Experiment 3 of the original work by Gratton et al. (1992) suggests that, like response conflict, arbitrary cues also may trigger the implementation of top-down control.

Whatever the explanation for the findings of Mayr et al. (2003), the present data indicate that those findings may not widely generalize a consideration that limits their theoretical implications. Instead, taken together with converging evidence from other domains, the present findings support the view that conflict monitoring may play a role in the modulation of ongoing task performance.

A final consideration is necessary to address some differences in the adjustments subsequent to response conflict between the two experiments. The data from Experiment 2 can be understood in terms of a twofold response to conflict involving (1) an increased focus on the target item and (2) a raising of response threshold. Experiment 2 further suggests that the latter adjustment can effectively mask the effect of focusing on incompatible trials. Interestingly, data from the accuracy block in Experiment 1 do not reflect the adjustment of response threshold (which would result in between-trials speed–

accuracy changes). This may be explained by the response deadline, which limits the range in which response threshold can be varied. Thus, the present study suggests that task context may influence the form of adjustments resulting from response conflict.

REFERENCES

BOTVINICK, M. M., BRAVER, T. S., BARCH, D. M., CARTER, C. S., & COHEN, J. D. (2001). Conflict monitoring and cognitive control. *Psychological Review*, **108**, 624-652.

BOTVINICK, M. [M.], NYSTROM, L. E., FISSELL, K., CARTER, C. S., & COHEN, J. D. (1999). Conflict monitoring versus selection-for-action in anterior cingulate cortex. *Nature*, **402**, 179-181.

CARTER, C. S., BRAVER, T. S., BARCH, D. M., BOTVINICK, M. M., NOLL, D., & COHEN, J. D. (1998). Anterior cingulate cortex, error detection, and the online monitoring of performance. *Science*, **280**, 747-749.

CORBALLIS, P. M., & GRATTON, G. (2003). Independent control of processing strategies for different locations in the visual field. *Biological Psychology*, **64**, 191-209.

ERIKSEN, B. A., & ERIKSEN, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception & Psychophysics*, **16**, 143-149.

GRATTON, G., COLES, M. G. H., & DONCHIN, E. (1992). Optimizing the use of information: Strategic control of activation of responses. *Journal of Experimental Psychology: General*, **121**, 480-506.

KERNS, J. G., COHEN, J. D., MACDONALD, A. W., III, CHO, R. Y., STENGER, V. A., & CARTER, C. S. (2004). Anterior cingulate conflict monitoring and adjustments in control. *Science*, **303**, 1023-1026.

LOFTUS, G. R., & MASSON, M. E. J. (1994). Using confidence intervals in within-subject designs. *Psychonomic Bulletin & Review*, **1**, 476-490.

MACDONALD, A. W., III, COHEN, J. D., STENGER, V. A., & CARTER, C. S. (2000). Dissociating the role of the dorsolateral prefrontal and anterior cingulate cortex in cognitive control. *Science*, **288**, 1835-1838.

MAYR, U., AWH, E., & LAUREY, P. (2003). Conflict adaptation effects in the absence of executive control. *Nature Neuroscience*, **6**, 450-452.

ROGERS, R. D., & MONSELL, S. (1995). Costs of a predictable switch between simple cognitive tasks. *Journal of Experimental Psychology: General*, **124**, 207-231.

SOHN, M. H., & CARTER, C. S. (2003). *Conflict adaptation is independent of stimulus repetition: Evidence for the conflict monitoring model*. Manuscript submitted for publication.

STADLER, M. A., & HOGAN, M. E. (1996). Varieties of positive and negative priming. *Psychonomic Bulletin & Review*, **3**, 87-90.

ULLSPERGER, M., & SZYMANOWSKI, F. (2004). ERP correlates of error relevance. In M. Ullsperger & M. Falkenstein (Eds.), *Errors, conflicts, and the brain: Current opinions on performance monitoring* (pp. 171-177). Leipzig: MPI for Human Cognitive and Brain Sciences.

ULLSPERGER, M., & VON CRAMON, D. Y. (2001). Subprocesses of performance monitoring: A dissociation of error processing and response competition revealed by event-related fMRI and ERPs. *NeuroImage*, **14**, 1387-1401.

NOTES

1. Gratton et al. (1992) did not report separate statistical analyses for target alternation trials alone, however.

2. The trend toward overall slower responses following incompatible trials may account for the fact that RTs for incompatible trials did not differ on the basis of preceding trial type [$t(23) = 0.58$, $p > .05$]. The difference for compatible trials was found to be statistically significant [$t(23) = 3.05$, $p < .01$].