Psychophysiology, 41 (2004), 157–160. Blackwell Publishing Inc. Printed in the USA.

DOI: 10.1046/j.1469-8986.2003.00128.x

BRIEF REPORT

Stimulus modality, perceptual overlap, and the go/no-go N2

SANDER NIEUWENHUIS,[a] NICK YEUNG,[a] AND JONATHAN D. COHEN[a,b]

[a] Department of Psychology, Princeton University, Princeton, New Jersey, USA
[b] Center for the Study of Brain, Mind, and Behavior, Princeton University, Princeton, New Jersey, USA

Abstract

Stimuli that elicit a prepotent but incorrect response are typically associated with an enhanced electrophysiological N2 that is thought to index the operation of a control process such as inhibition or conflict detection. However, recent studies reporting the absence of the N2 modulation in go/no-go tasks involving auditory stimuli challenge this view: It is not clear why inhibition or conflict detection should be sensitive to the modality of the stimulus. Here we present electrophysiological data from a go/no-go task suggesting that the relative size of the N2 modulation in visual and auditory tasks depends on the perceptual overlap between the go and no-go stimuli. Stimuli that looked similar but sounded different were associated with a typical visual N2 modulation and the absence of an auditory N2 modulation, consistent with previous findings. However, when we increased the perceptual overlap between the auditory stimuli, a large no-go N2 was observed. These findings are discussed in terms of existing hypotheses of the N2, and clarify why previous studies have not found an N2 modulation in auditory go/no-go tasks.

Descriptors: No-go, Stimulus modality, Response conflict, N2, Monitoring, Inhibition

There has been increasing interest in electrophysiological correlates of cognitive control processes, the high-level processes that monitor and regulate basic information processing. An important example concerns the N2, a negative event-related brain potential (ERP) component with a frontocentral scalp distribution that peaks around 300 ms after stimulus presentation. The N2 is often enhanced under conditions in which there is a tendency to make a prepotent but incorrect response (e.g., Kopp, Rist, & Mattler, 1996; Pfefferbaum, Ford, Weller, & Kopell, 1985; van Boxtel, van der Molen, Jennings, & Brunia, 2001). An example of a laboratory task involving such conditions is the go/no-go task, in which partipants are instructed to give a speeded, simple response to frequent target ("go") stimuli and to withhold a response to infrequent distractor ("no-go") stimuli. No-go stimuli are associated with a larger N2 than go stimuli, especially when the go response is primed (Eimer, 1993; Jodo & Kayama, 1992; Kok, 1986).

Although there is debate over whether the N2 reflects the inhibition of the prepotent (e.g., go) response (e.g., Falkenstein, Hoormann, & Hohnsbein, 1999; Kok, 1986; Kopp et al., 1996) or the detection of response conflict (Nieuwenhuis, Yeung, Van den Wildenberg, & Ridderinkhof, 2003; Van Veen & Carter, 2002; Yeung, Botvinick, & Cohen, 2003), there appears to be consensus that the N2 is a marker of a general control process that operates in a variety of situations. However, this view has been challenged by a number of studies reporting that the typical N2 amplitude enhancement for no-go stimuli is seen for visual stimuli, but not for auditory stimuli (Falkenstein et al., 1999; Falkenstein, Hoormann, & Hohnsbein, 2002; Falkenstein, Koshlykova, Kiroj, Hoormann, & Hohnsbein, 1995). The finding that the N2 is modulated by stimulus input modality seems hard to reconcile with hypotheses that associate this component with a general (modality-independent) control process. To account for their findings, Falkenstein and colleagues (e.g., Falkenstein et al., 1999) suggested that the N2 may reflect "a modality-specific … inhibition process" (p. 289). However, this account is somewhat unsatisfactory, because it is unclear why there should be separate response inhibition processes for visual and auditory stimuli.

The experiment reported here was designed to test an alternative hypothesis of why the no-go N2 appears to be sensitive to input modality. In each of their studies, Falkenstein and colleagues used letter stimuli, presented either visually or as speech. For instance, in two studies (Falkenstein et al., 1995, 1999) the letters "F" and "J" served as go and no-go stimuli. We noticed that these letters are easy to discriminate when spoken (in English and in German), but less so when presented visually, due

---

The research of S. Nieuwenhuis was supported by the Netherlands Organization for Scientific Research (NWO). We thank Chris Asplund and Rebecca Hackell for technical assistance, and Richard Ridderinkhof for comments on an earlier draft of the manuscript.

Address reprint requests to: Sander Nieuwenhuis, who is now at the Department of Cognitive Psychology, Vrije Universiteit, Van der Boechorststraat 1, 1081 BT Amsterdam, The Netherlands. E-mail: stn20@dds.nl.

to a number of visual features that they have in common. Importantly, the degree of perceptual overlap between the stimuli may modulate the erroneous activation of the go response on no-go trials, yielding more interference when the stimuli are similar. The pattern of behavioral data in the studies of Falkenstein and colleagues was consistent with this conjecture: Auditory stimuli were generally associated with lower false alarm rates and faster go RTs than visual stimuli. According to the conflict monitoring hypothesis, differential activation of the go response should affect the degree of response conflict on no-go trials (Jones, Cho, Nystrom, Cohen, & Braver, 2002), and hence N2 amplitude. The inhibition hypothesis makes a similar prediction. Thus, we hypothesized that the previously reported modality-related differences in no-go N2 amplitude are not due to differences in stimulus modality per se, but instead to differences in perceptual overlap of the stimuli when presented in the visual and auditory modality.

At first glance, this hypothesis would seem at odds with a report by Kiefer and colleagues (Kiefer, Marzinzik, Weisbrod, Scherg, & Spitzer, 1998), who found no frontocentral N2 modulation despite similar auditory go and no-go stimuli. However, in this study, participants were not instructed to emphasize response speed over accuracy, as is evidenced by the long mean reaction times (RTs; >500 ms) and low false alarm rates ($<5\%$) in each of the task conditions (M. Kiefer, pers. comm.). Speed pressure, which modulates the bias toward the go response, is known to be an important determinant of the size of the N2 modulation (e.g., Jodo & Kayama, 1992).

In the present experiment, participants were presented, in separate trial blocks, with visual and auditory go and no-go stimuli. In each block, participants were required to suppress a response (or to respond) to the letter "F" in the context of responding (or suppressing a response) to another letter. We were interested in how the go and no-go ERPs associated with the letter "F" varied as a function of stimulus modality and context letter. In one task condition, the context letter ("T") looked similar but sounded different to the "F" stimulus. In this condition, we expected to replicate Falkenstein's (e.g., Falkenstein et al., 1999) findings: a substantial no-go N2 in the visual condition and a strongly reduced no-go N2 in the auditory condition. In a second task condition the context letter ("S") looked different but sounded similar to the "F" stimulus. Importantly, in this condition, we expected to find the opposite pattern of results: a substantial N2 enhancement to auditory no-go stimuli and a reduced visual no-go N2. To ensure that we observed robust N2 components on no-go trials, we imposed a stringent response deadline on go trials, and presented no-go stimuli with a low frequency (cf. Nieuwenhuis et al., 2003).

Method

Participants

Twelve young adults (6 women, *M* age = 23.5 years) participated for payment ($20).

Stimuli

Visual stimuli were presented in white against a black background on a computer screen placed at a distance of 150 cm from the participant. The stimuli were the letters "F," "S," and "T," presented in a Times New Roman font. The "F" was slightly modified to resemble the "T" more. The letters subtended 0.6° and were displayed just above a central fixation dot (0.3°) that remained on the screen throughout the experiment. Auditory stimuli were the letters F, S, and T, similar in pitch and volume, and spoken by a female native English speaker. Separate in-ear noise-insulating headphones were employed. Before the experimental session, participants adjusted the loudness of the speech stimuli to match the brightness of the visual stimuli, using a standard subjective intensity-matching procedure (cf. Falkenstein et al., 1999).

Design and Procedure

The experiment consisted of four sequences of four blocks of 100 trials each. Stimuli were presented in one modality in sequences 1 and 4, and in the other modality in sequences 2 and 3. Response finger (left or right index finger) was changed halfway through the experiment. The order of these two factors was varied orthogonally across participants. The identity of go and no-go stimuli (F&S, S&F, F&T, or T&F, respectively) was varied across the four blocks in each sequence. The order of the stimulus identities was counterbalanced over sequences and participants using a Latin square design. Each block contained 25 no-go trials and 75 go trials presented in a pseudorandom order.

Each trial began with the presentation of a stimulus for 200 ms. The interval separating successive stimuli was one of five equiprobable durations (1.1, 1.3, 1.5, 1.7, or 1.9 s). Participants were required to make a button press within 400 ms after the onset of each go stimulus. If a response exceeded the deadline, an auditory tone (100 ms, 800 Hz) presented 1 s after the stimulus informed the participant that she or he should try to respond faster on the next trial. Participants were informed about the input modality and the identities of the go and no-go stimuli in advance of each block. Before each sequence of blocks, they received 20 practice trials. Participants were instructed to press a response button before the deadline on go trials while avoiding button presses on no-go trials. They were given feedback at the end of each block showing their mean go RT and false alarm rate in that block. There were 5-min breaks between sequences of blocks.

Psychophysiological Recording and Data Analysis

EEG recordings were taken from 18 Ag/AgCl electrodes placed in a fabric cap (Electro-Cap International, Inc.), referenced to the left earlobe: FP1, FP2, AFz, F3, Fz, F4, FC3, FCz, FC4, C3, Cz, C4, CP3, CPz, CP4, P3, Pz, and P4. The electrooculogram (EOG) was recorded from electrodes placed above and below the left eye, and from electrodes placed on the outer canthi of each eye. The ground electrode was placed on the chin. All electrode impedances were kept below 30 kΩ. The EEG signals were amplified (Sensorium Inc. EPA-6; bandpass filter 0.1–300 Hz), and digitized at 600 Hz.

Single-trial epochs were extracted off-line. The EEG data were re-referenced to linked-earlobe electrodes. The EMCP method (Gratton, Coles, & Donchin, 1983) was used to correct for EOG artifacts and to discard trials with amplifier saturation. For each participant and each condition, the EEG epochs were averaged with respect to stimulus onset. Before subsequent analyses, the resulting ERP waveforms were lowpass filtered ($<$12 Hz) using a third-order digital Butterworth filter. N2 amplitude was computed on the basis of the signals obtained from Cz, where the component was largest. We selected the first negative peak in the window 200–400 ms following the stimulus. We then identified the positive peak preceding this negative peak. N2 amplitude was defined as the amplitude of the negative peak minus the amplitude of the positive peak. Behavioral data and

Table 1. *Behavioral Performance Data as a Function of Task Condition*

	Go trials		No-go trials	
Task condition	RT (ms)	% misses	FA RT (ms)	% FA
Visual/Context T	339 (10)	1.6 (0.5)	295 (12)	39.0 (3.2)
Auditory/Context T	285 (10)	6.9 (1.2)	304 (12)	15.5 (3.2)
Visual/Context S	325 (5)	1.3 (0.4)	285 (8)	28.3 (2.8)
Auditory/Context S	330 (9)	4.8 (0.8)	275 (14)	32.5 (4.8)

Note: Values are means with standard errors of the mean in parentheses. FA: false alarms. Behavioral data are based on trials in which the presented go or no-go stimulus was an "F."

N2 amplitudes were analyzed using repeated measures analyses of variance with factors context letter (T or S), modality (visual or auditory), and trial type (go or no-go; only for the N2 analysis).

Results

The analyses of behavioral and N2 data were confined to the trials in which the presented go or no-go stimulus was an "F." This yielded an unconfounded estimate of the effects of context letter. Similar results were obtained for the other two letters.

Behavioral Results

Table 1 shows behavioral data for each task condition. RT to go stimuli was reliably affected by modality, $F(1,11) = 11.1$, $p = .007$, and context letter, $F(1,11) = 18.6$, $p = .001$. Importantly, these two main effects were qualified by a significant interaction effect, $F(1,11) = 86.8$, $p < .001$. Modality had opposite numerical effects for the two context letters: When the context letter looked similar but sounded different ("T"), RTs were faster for auditory stimuli than for visual stimuli. In contrast, when the context letter looked different but sounded similar ("S"), RTs were faster for visual stimuli than for auditory stimuli. The simple main effect of modality was reliable for context letter "T," $F(1,11) = 38.2$, $p < .001$, but not for context letter "S," $F(1,11) < 1$.

Although false alarm rate was generally higher in the visual than in the auditory conditions, $F(1,11) = 23.3$, $p = .001$, the modality effect went in opposite directions for the different context letters, as reflected in a significant interaction effect, $F(1,11) = 33.9$, $p < .001$. When the context letter looked similar but sounded different ("T"), participants made fewer false alarms to auditory stimuli than to visual stimuli. In contrast, when the context letter looked different but sounded similar ("S"), participants made fewer false alarms to visual stimuli than to auditory stimuli. The simple main effect of modality was reliable for context letter "T," $F(1,11) = 90.1$, $p < .001$, but not for context letter "S," $F(1,11) = 1.3$, $p = .28$.

N2 Amplitude

Figure 1 shows the ERP waveforms associated with correct go and no-go trials in each of the four conditions of the experimental design. The N2 is clearly visible as a negative deflection peaking roughly 300 ms after stimulus onset. N2 amplitude was generally larger on no-go trials ($M = -9.7\,\mu$V) than on go trials ($M = -6.7\,\mu$V), $F(1,11) = 63.2$, $p < .001$. This N2 modulation was reliably affected by modality, $F(1,11) = 44.1$, $p = .004$, and context letter, $F(1,11) = 23.9$, $p < .001$. The effect of modality on the size of the N2 modulation was significantly different for the two context letters, as evidenced by a reliable three-way interaction effect, $F(1,11) = 10.8$, $p = .007$. When the context letter looked similar but sounded different ("T"), the N2 modulation was larger in the visual condition ($M = -4.4\,\mu$V) than in the auditory condition ($M = +1.7\,\mu$V). In contrast, when the context letter looked different but sounded similar ("S"), the N2 modulation was larger in the auditory condition

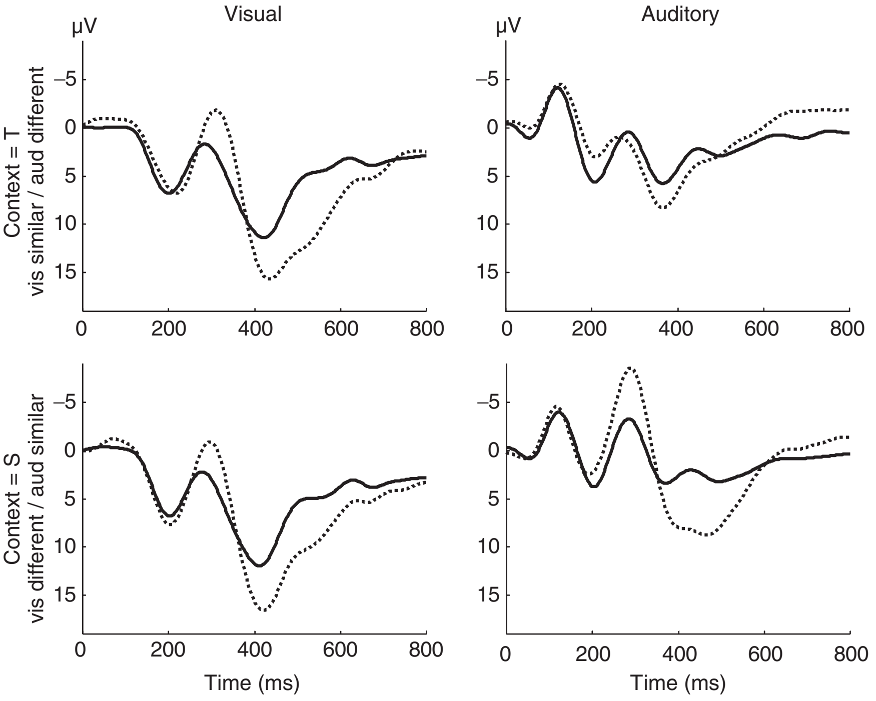

Figure 1. Stimulus-locked ERPs at electrode Cz for correct go and no-go trials, separated by task condition. Time = 0 ms indicates stimulus onset. A 200-ms prestimulus baseline was subtracted from each ERP. Vis: visually; Aud: auditorily.

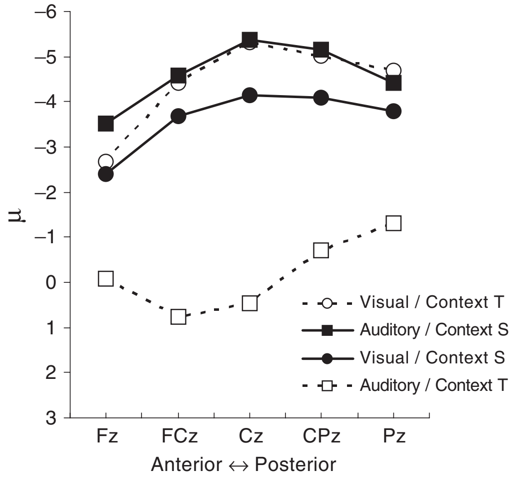

Figure 2. Peak amplitude at five midline electrodes of the N2 modulation in each task condition. For each condition and electrode, the N2 modulation was computed as the most negative value of the ERP difference wave (no-go minus go) in the window 275–325 ms following the stimulus.

($M = -4.9\,\mu$V) than in the visual condition ($M = -4.3\,\mu$V). The simple interaction between modality and trial-type was significant for context letter "T," $F(1,11) = 16.2$, $p = .002$, but not for context letter "S," $F(1,11) < 1$. Figure 2 shows that the N2 modulation exhibited a similar midline scalp topography in the three conditions in which it was observed.

Discussion

Stimuli that elicit a prepotent but incorrect response are often associated with an amplitude modulation of the electrophysiological N2. This N2 component is commonly believed to reflect the operation of a general evaluative or executive control mechanism (for a discussion, see Nieuwenhuis et al., 2003). However, this interpretation has been complicated by a series of observations suggesting that the modulation is absent in auditory go/no-go tasks (Falkenstein et al., 1995, 1999, 2002). We hypothesized that the absence of a no-go N2 enhancement in these earlier studies may reflect the stimuli used: In general, the auditory stimuli seemed more easily discriminable than did the visual stimuli, which may have considerably weakened the tendency to activate the go response on no-go trials. In that case, the conflict-monitoring hypothesis would predict a reduced N2 on correct no-go trials (cf. Nieuwenhuis et al., 2003). The inhibition hypothesis would make the same prediction: There is less need to inhibit the go response (cf. Jodo & Kayama, 1992).

The present experiment produced results consistent with the perceptual-overlap hypothesis. In one condition (context letter "T"), the go and no-go stimuli were more easily discriminable when presented auditorily than when presented visually: Auditory stimuli were associated with fewer false alarms and faster go RTs than visual stimuli. In this condition, we replicated the results reported by Falkenstein and colleagues (Falkenstein et al., 1995, 1999, 2002): a typical enhancement of the no-go N2 following visual stimuli but not following auditory stimuli. In the other condition (context letter "S"), discrimination of the auditory stimuli was at least as difficult as discrimination of the visual stimuli. Here we found a substantial modulation of the auditory no-go N2. This modulation was larger than that of the visual no-go N2, though not reliably so. We expect that the visual N2 modulation would have been smaller had we succeeded in choosing more easily discriminable visual stimuli; performance for these stimuli was only slightly worse than for the auditory stimuli. Nevertheless, our results show that the presence or absence of a no-go N2 modulation in auditory tasks is dependent on the perceptual overlap between the stimuli, and hence is under experimental control. These findings are consistent with the view that the N2 is an electrophysiological marker of a general cognitive control process.

REFERENCES

Eimer, M. (1993). Effects of attention and stimulus probability on ERPs in a Go/Nogo task. *Biological Psychology*, *35*, 123–138.

Falkenstein, M., Hoormann, J., & Hohnsbein, J. (1999). ERP components in Go/Nogo tasks and their relation to inhibition. *Acta Psychologica*, *101*, 267–291.

Falkenstein, M., Hoormann, J., & Hohnsbein, J. (2002). Inhibition-related ERP components: Variations with modality, age, and time-on-task. *Journal of Psychophysiology*, *16*, 167–175.

Falkenstein, M., Koshlykova, N. A., Kiroj, V. N., Hoormann, J., & Hohnsbein, J. (1995). Late ERP components in visual and auditory Go/Nogo tasks. *Electroencephalography and Clinical Neurophysiology*, *96*, 36–43.

Gratton, G., Coles, M. G., & Donchin, E. (1983). A new method for off-line removal of ocular artifact. *Electroencephalography and Clinical Neurophysiology*, *55*, 468–484.

Jodo, E., & Kayama, Y. (1992). Relation of a negative ERP component to response inhibition in a go/no-go task. *Electroencephalography and Clinical Neurophysiology*, *82*, 477–482.

Jones, A. D., Cho, R., Nystrom, L. E., Cohen, J. D., & Braver, T. S. (2002). A computational model of anterior cingulate function in speeded response tasks: Effects of frequency, sequence, and conflict. *Cognitive, Affective, & Behavioral Neuroscience*, *2*, 300–317.

Kiefer, M., Marzinzik, F., Weisbrod, M., Scherg, M., & Spitzer, M. (1998). The time course of brain activations during response inhibition: Evidence from event-related potentials in a go/no go task. *Neuroreport*, *9*, 765–770.

Kok, A. (1986). Effects of degradation of visual stimulation on components of the event-related potential (ERP) in Go/No-Go reaction tasks. *Biological Psychology*, *23*, 21–38.

Kopp, B., Rist, F., & Mattler, U. (1996). N200 in the flanker task as a neurobehavioral tool for investigating executive control. *Psychophysiology*, *33*, 282–294.

Nieuwenhuis, S., Yeung, N., Van den Wildenberg, W., & Ridderinkhof, K. R. (2003). Electrophysiological correlates of anterior cingulate function in a Go/NoGo task: Effects of response conflict and trial-type frequency. *Cognitive, Affective, & Behavioral Neuroscience*, *3*, 17–26.

Pfefferbaum, A., Ford, J. M., Weller, B. J., & Kopell, B. S. (1985). ERPs to response production and inhibition. *Electroencephalography and Clinical Neurophysiology*, *60*, 423–434.

van Boxtel, G. J., van der Molen, M. W., Jennings, J. R., & Brunia, C. H. (2001). A psychophysiological analysis of inhibitory motor control in the stop-signal paradigm. *Biological Psychology*, *58*, 229–262.

Van Veen, V., & Carter, C. S. (2002). The timing of action monitoring processes in anterior cingulate cortex. *Journal of Cognitive Neuroscience*, *14*, 593–602.

Yeung, N., Botvinick, M. M., & Cohen, J. D. (2003). The neural basis of error detection: Conflict monitoring and the error-related negativity. Manuscript submitted for publication.

(Received February 14, 2003; Accepted June 23, 2003)