Psychonomic Bulletin & Review
2008, 15 (6), 1117-1121
doi:10.3758/PBR.15.6.1117

He said, she said: Episodic retrieval induces conflict adaptation in an auditory Stroop task

Michiel M. Spapé and Bernhard Hommel
Leiden University, Leiden, The Netherlands

People respond more slowly if an irrelevant feature of a target stimulus is incompatible with the relevant feature or the correct response. Such compatibility effects are often reduced in trials following an incompatible trial, which has been taken to reflect increased cognitive control. This pattern holds only if two trials share some similarities, however, suggesting that it may be modulated by the episodic context. To look into this possibility, we had participants respond to high- or low-pitched tones by saying "high" or "low," respectively, and ignore the simultaneously presented auditory word "high" or "low." As expected, performance was impaired if the heard word was incompatible with the required response, and this Stroop-like effect was reduced after incompatible trials. This sequential modulation was observed, however, only if the voice in the two successive trials was the same, whereas no modulation was obtained when the speaker changed. The results suggest that sequential modulations are due to the automatic retrieval of episodic event representations that integrate stimuli, actions, and situational and task-specific control information, so that later reactivation of some elements of a given representation tends to retrieve the other elements as well.

Human performance is very sensitive to conflict between stimuli and the responses they activate. Reactions are slower and less accurate if the features or elements of a stimulus point to different actions, as indicated by the Stroop effect (Stroop, 1935), the Simon effect (Simon & Rudell, 1967), and the flanker effect (Eriksen & Eriksen, 1974). Consider an example from the vocal Stroop task we used in the present study: Vocally judging high- and low-pitched tones as "high" and "low" goes faster if they are accompanied by congruent distractors (e.g., the words "high" and "low," respectively) than if accompanied by incongruent distractors (the words "low" and "high"). Recent studies have shown, however, that people can adapt to conflict-inducing circumstances, probably by using the occurrence and degree of conflict between cognitive representations to fine-tune control processes so as to reduce or even eliminate conflict in the future (Botvinick, Nystrom, Fissell, Carter, & Cohen, 1999). Consistent with this suggestion are observations of systematic sequential changes of conflict-induced effects: The Stroop effect (Kerns et al., 2004), the Simon effect (Stürmer, Leuthold, Soetens, Schröter, & Sommer, 2002), and the flanker effect (Gratton, Coles, & Donchin, 1992) are all smaller, absent, or negative after a conflict trial, and larger after a nonconflict trial. Hence, one is likely to suffer less from an incongruent target–distractor combination (say, a low tone and the word "high") after having just experienced another incongruent combination (say, a high tone and the word "low") than after a congruent combination. Apparently, then, registering conflict in a given trial increases the amount of cognitive control exerted, so that conflict can be prevented, or at least reduced, in the subsequent trial.

Even though this cognitive control account fits much of the obtained data patterns, recent observations show that sequence modulations of conflict effects reflect more than the workings of executive control. Several authors have pointed out that repetitions and alternations of conflict trials are commonly confounded with repetitions and alternations of relevant and irrelevant stimulus and response features and, most critically, of the relations among them (e.g., Hommel, Proctor, & Vu, 2004; Mayr, Awh, & Laurey, 2003; Nieuwenhuis et al., 2006; Wendt, Kluwe, & Peters, 2006). Consider, for instance, the incongruent combination of a high tone and the distractor "low." The fact that people do better if they have just been presented with another incongruent combination (low tone and "high," say) than with a congruent combination (high tone and "high," say) may have to do with the repetition of incongruence, but it may also be due to a kind of negative priming effect: Note that the "congruent→incongruent" sequence implies a partial but incomplete feature overlap between the two succeeding stimulus–response episodes, because the tone and the response repeat but the distractor changes. In contrast, the "incongruent→incongruent" sequence does not entail any such overlap. Partial overlap of succeeding stimulus–response episodes is known to produce performance costs even in the absence of any target–distractor incongruence, presumably because feature repetitions retrieve the previous episode and thus in-

M. M. Spapé, mspape@fsw.leidenuniv.nl

 Copyright 2008 Psychonomic Society, Inc.

duce conflict if previous and present episodes do not fully match (Hommel, 1998, 2004). This would be the case if congruent conditions followed incongruent conditions or vice versa, but not if congruent or incongruent conditions followed each other—an outcome pattern that may only look like control-induced adaptation effects.[1]

Recent studies sought to disentangle episodic and control effects and provided evidence that sequential modulations of conflict effects may actually reflect *both* episodic retrieval (e.g., Hommel et al., 2004; Wendt et al., 2006) *and* control-induced effects (e.g., Akçay & Hazeltine, 2007; Wühr, 2005). Proponents of episodic and control approaches have tended to treat these demonstrations as existence proofs, as it were, to justify further investigation of episodic and control effects in isolation. However, here we would like to argue that episodic processes and control operations may not just happen to co-occur in the mentioned conflict tasks, but may be intrinsically related in a theoretically interesting way. In particular, it may be that episodic representation of events—event files in the sense of Hommel (1998, 2004)—integrate not only stimulus and response codes, but also task and control parameters, including the weights of associations between stimulus and response features that control theorists assume to produce the control-related part of sequential modulations (e.g., Botvinick et al., 1999).

Some evidence for this possibility has been obtained in task-switching studies: Processing a stimulus previously encountered in another task interferes with switching to a new task, suggesting that the stimulus induced the automatic retrieval of a representation of the task it accompanied (Waszak, Hommel, & Allport, 2003). If we assume that a stored task representation comprises stimulus–response associations, it makes sense to assume that the strengths of these associations can be stored with them—which implies that the outcomes of the hypothesized control operations may be encoded in episodic event representations and later retrieved with them. This consideration is consistent with recent observations from sequential studies. Some authors have reported that conflict adaptation after a task switch is observed only if there is at least some overlap in terms of stimulus or response features between the two tasks (e.g., Kiesel, Kunde, & Hoffmann, 2006). In the absence of a certain degree of similarity between tasks, Notebaert and Verguts (2008) found no remaining effect of one conflicting trial on the next, prompting them to state that conflict monitoring acts *locally* (see also Blais, Robidoux, Risko, & Besner, 2007). Indeed, this is what an episodic integration approach would suggest: Stimulus–response episodes belonging to the same or to similar tasks overlap the contextual and task information they contain, so that they are more likely to be retrieved in the same context.

The degree to which similarity between two episodes affects cognitive control may be related to considerations proposed by the event segmentation theory of Zacks and colleagues (Zacks & Swallow, 2007; Zacks & Tversky, 2001). It assumes that people continuously predict future perceptions on the basis of the flow of current perceptions but will segment the stream into a new event if the error between prediction and perception becomes too large—that is, if critical features change. In other words, if two external events are too different from one another to fit comfortably into one event representation, a new event boundary is created to make sense of ongoing activity. Although what exactly constitutes "too large" remains fuzzy, there is some evidence that people are likely to create a new event boundary during changes in time, space, and causal contingency—for example, if the protagonist in a story changes (Kurby & Zacks, 2008). This may suggest that encountering a significant change in a context variable (such as the task) resets the system, thereby eliminating the aftereffects of conflict-induced adjustments—a line of thought that would fit the idea that task switching involves the inhibition of the previous task (Mayr & Keele, 2000) or the flushing of control systems (Logan & Gordon, 2001).

To explore the possibility that conflict-induced effects are more likely in situations where two successive events are similar, we manipulated a feature that has no obvious relation to action control operations and does not call for a task switch, but nevertheless constitutes a salient change in the episodic context. As has already been mentioned, we presented participants with a vocal Stroop-like task where they were to respond to a high- or low-pitched tone while ignoring a randomly varying congruent or incongruent word ("high" or "low"). The episodic manipulation of interest had a direct relationship to the voice speaking the word, which either changed or remained the same from one trial to another. Voice can be discriminated very rapidly and induces the automatic classification of a speaker; this has been found to channel the integration of information from the spoken message (Van Berkum, Van den Brink, Tesink, Kos, & Hagoort, 2008). If so, voice changes should be particularly suited to signal event boundaries and the closing of open event representations and the beginning of a new event. In other words, changing the voice (even though entirely irrelevant for the task) can be expected to alter the context and/or trigger event segmentation processes, rendering a sequence of trials to be less similar to sequences in which the voice was repeated. If so, conflict-induced sequential effects may be smaller or even absent after a voice change.

METHOD

Participants

Seven female and 7 male students from Leiden University voluntarily participated in this experiment for a small fee or course credits. Data from 1 more female participant did not enter analysis, due to an error rate of more than 50%.

Stimuli

Participants responded to sinusoidal tones of 550 or 1050 Hz by vocally responding "low" and "high" (in Dutch), respectively. The irrelevant stimulus was the word "low" or "high" (in Dutch), spoken by a male voice, a female voice, or the participant (recorded and edited before the experiment started).[2] Voices were recorded and later presented using a Samson microphone attached to the Behringer BCA2000 audio interface. To present the samples at similar psychoacoustic levels, they were normalized to maximum level, trimmed to 350 msec, compressed at a rate of 2:1 above −15 dB, and then again normalized to −6 dB level, using Syntrillium CoolEdit. A sinusoidal tone of 550 or 1050 Hz of −9 dB was generated and digitally mixed into the samples. Finally, the tone and voice were reshaped with a linear slope on the first and final 10 msec, respectively, to compensate for onset and offset clicking sounds, after which the samples were normalized again to maximum level.

Procedure

Stimuli were presented with AKG K-55 headphones and timed using E-Prime 1.1.3, under Windows XP SP2, running on a 1.8-GHz Dell Latitude laptop. A Psychology Software Tools voice key setup recorded vocal reactions, which were displayed on a computer screen that was monitored for accuracy and false alarms by two independent observers. White noise was played during the entire experiment to mask external and extra-experimental sound.

The experiment consisted of pairs of trials, in which the first member of each pair (stimulus S1 and response R1) was considered the prime, and the second member (S2 and R2) the probe, with the congruence relationship between prime and probe being the interesting manipulation. Each pair of trials started by playing white noise for 150 msec, a preparation signal that was followed by 1,000 msec of silence. The first stimulus (S1) was then played for 350 msec. Participants were required to respond to the high- or low-pitched tone by saying "high" or "low" into the microphone (R1), ignoring the voice they heard along with the tone. After they responded, another 400 msec passed before the onset of the second stimulus (S2), which was also played for 350 msec and was responded to by participants' vocally judging its pitch as "high" or "low" (R2). The pair of trials was concluded with a 1,000-msec silent interval—unless the participant had reacted incorrectly, in which case verbal feedback was given. The first 20 pairs of trials were considered training, during which the participant received extra feedback concerning speed and accuracy of performance, and after 144 pairs, a 5-min break was introduced, followed by another 144 pairs of trials.

The experiment used a three-factor repeated measures design with the factors S1–R1 congruency (where combinations of high pitch and the distractor word "high" and of low pitch and the word "low" were considered congruent, and the other combinations incongruent), S2–R2 congruency, and voice repetition. The combinations of the type of voice, the irrelevant word, and the pitch of the tone for both S1–R1 and S2–R2 were randomized but balanced across design cells.

RESULTS

In repeated measures ANOVAs of responses to S1—reaction times (RTs) and error rates (ERs)—with the factors of word ("high" or "low"), response (to a high- or low-pitched tone), and voice, responding with the word "high" was found to occur significantly faster than responding with the word "low" [$F(1,13) = 14.87, p < .003, \eta^2 = .53$], although not significantly more accurate ($p > .2$). The distracting word had no significant effect on RT ($p > .4$) or ER ($p > .1$). A congruency effect was revealed by interactions between word and response on both RTs [$F(1,13) = 25.59, p < .0003, \eta^2 = .66$] and ERs [$F(1,13) = 7.89, p < .02, \eta^2 = .38$], with incongruent words resulting in reactions 65 msec slower, with 6% more errors. However, this effect was not modulated by voice (ps $> .2$), nor did voice interact with other variables (ps $> .1$).

An overview of the S2–R2 results is provided in Figure 1. With voice repetition conditions (left panel), a much greater S2–R2 congruency effect (see the distance between dotted and straight lines) was found after congruent S1–R1 conditions (left part of the left panel) than after incongruent S1–R1 conditions (right part of the left panel). This was not the case for voice alternation conditions (right panel): Here, the congruency effect of S2–R2 was independent of S1–R1 congruence.

The ANOVAs—with S1 congruency, S2 congruency, and voice repetition or alternation as factors—confirmed this pattern. Although S1 congruency had no effect (ps $> .1$), S2 congruency affected RTs [$F(1,13) = 21.70, p < .0005, \eta^2 = .63$] and ERs [$F(1,13) = 14.41, p < .003, \eta^2 = .53$]: Responses were 54 msec slower and 6% less

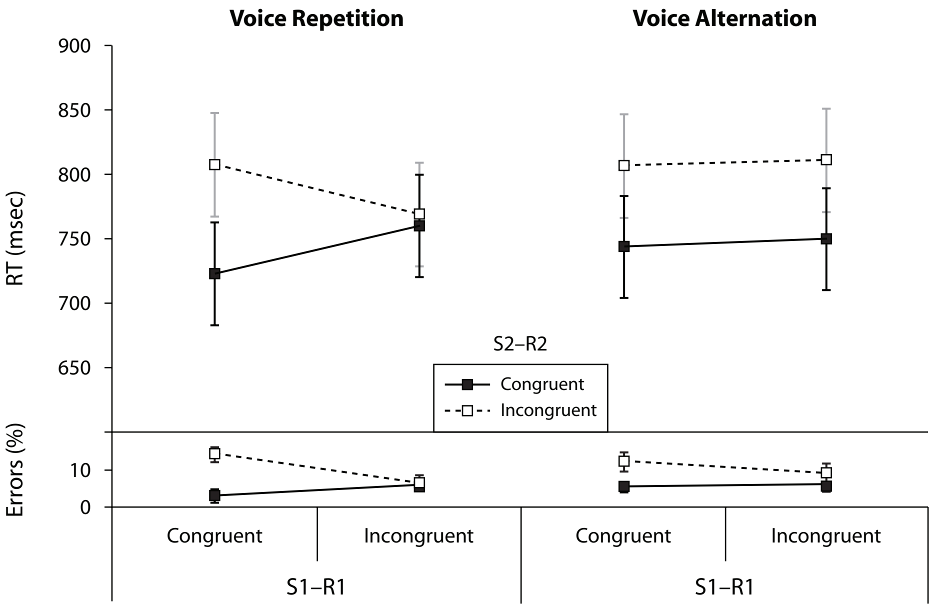

Figure 1. Reaction times (RTs) and errors for S2–R2 as a function of S2–R2 congruency, S1–R1 congruency, and the type of voice repetition between the two trials.

Table 1
Reaction Times (RTs) and Error Rates (ERs) for S2–R2 As a Function of S2–R2 Congruency, S1–R1 Congruency, and Voice Repetition or Alternation

		S2–R2				Effects			
		Congruent		Incongruent		Conflict		Conflict Adaptation	
Voice	S1–R1	RT	ER	RT	ER	RT	ER	RT	ER
Repeated	Congruent	723	3.2	807	14.2	85	11.0		
	Incongruent	760	6.2	769	6.7	9	0.5	76	10.5
Alternated	Congruent	744	5.1	807	12.1	64	7.0		
	Incongruent	750	5.3	811	9.0	60	3.7	4	3.4

Note—Effect sizes show the congruency effect (conflict) and how it is affected by congruency in $n-1$.

accurate in incongruent than in congruent conditions. Repeating the voice showed a trend toward significance in RT [$F(1,13) = 3.16, p = .099, \eta^2 = .20$], with a repeating voice giving a small 13-msec benefit.

More important were the interactions. As expected, the two-way interaction between S1–R1 congruency and S2–R2 congruency affected RTs [$F(1,13) = 9.30, p < .01, \eta^2 = .42$] and ERs [$F(1,13) = 7.94, p < .02, \eta^2 = .38$], with the S2–R2 congruency effect being 38 msec smaller after an incongruent trial than after a congruent one. This conflict-adaptation effect was significantly modulated by a three-way interaction involving voice repetition in RTs [$F(1,13) = 16.36, p < .002, \eta^2 = .56$], and nearly so in ERs [$F(1,13) = 4.66, p < .06, \eta^2 = .26$]; see Table 1. If the voice was repeated, the S2–R2 congruency effect was strongly dependent on S1–R1 congruency, the S2–R2 congruency effect decreasing by 76 msec (or 10.5% ER) after incongruent, not congruent, S1–R1 conditions [RT, $F(1,13) = 21.30, p < .0005, \eta^2 = .62$; ER, $F(1,13) = 14.56, p < .003, \eta^2 = .53$]. However, no such modulation was observed for voice alternations where the S2–R2 congruency effect decreased only 4 msec (or 3.4% ER) after incongruent rather than congruent S1–R1 conditions [RT, $F(1,13) < 1$; ER, $F(1,13) = 1.11, p > .3, \eta^2 = .08$].

DISCUSSION

Hearing a stimulus- and response-incongruent and irrelevant word impaired performance, in a manner very similar to that for the standard Stroop effect. In keeping with previous findings, this Stroop effect was considerably reduced after conflict trials, thus replicating Gratton et al.'s (1992) observation in the flanker paradigm and comparable findings of Kerns et al. (2004) and others using Stroop-type tasks. However, the previous trial affected the degree of conflict *only* if it was presented in the same voice. In contrast, no sequential effect was obtained if the voice changed. This observation goes beyond previous findings that sequential effects are affected by task switches or the similarity between sequential tasks by showing that sequential effects can be eliminated even within the same task and even with an entirely task-irrelevant similarity manipulation. We take that as a clear demonstration that the impact of conflict-induced operations is mediated by the episodic context, and we consider this contextual modulation to be why conflict-induced control acts "locally" (Blais et al., 2007; Notebaert & Verguts, 2008).

As we have pointed out, there are at least two possible ways in which the episodic context may interact with control processes or their aftereffects. First, each single stimulus–response episode may be integrated with the task context, including the just-modified weights of stimulus–response associations, and retrieved only in a sufficiently similar context, be it the task to be carried out (as with Kiesel et al., 2006) or the source of the information (as in the present study). Second, significant changes in the context may induce event segmentation, which in turn may lead to the inhibition of the representation of the previous event. Even though both interpretations would fit with our present observations, other findings seem to favor the first over the second. As was noted earlier, bindings between stimuli and task representations impair switching to another task. Given that they do so even if the acquisition of the binding and the task switch are separated by more than 100 trials (Waszak et al., 2003), this is unlikely to reflect transient inhibition but, rather, the retrieval of episodic traces. Along the same lines, conflict adaptation has been observed to operate across intervening trials. For instance, Wendt et al. (2006) had participants alternate between two different conflict tasks and found that the size of the conflict effect in the present trial (n) was considerably affected by the amount of conflict in the previous trial of the *same task* (i.e., trial $n-2$). This is inconsistent with the idea that context change or task switch would lead to or even require the inhibition of task representations. Accordingly, we attribute context effects to episodic retrieval. In particular, we suggest that episodic event representations integrate stimuli, actions, and situational and task-specific information, so that later reactivation of some elements of a given representation will tend to retrieve the other elements as well (Hommel, 2004). Even though more research is necessary to determine exactly what characteristics of an episode count as context, which aspects of context and task are considered for event integration, and how exactly the similarity between event codes is computed, it seems clear that event representations contain control parameters that are automatically retrieved under suitable conditions. This suggests that cognitive control and episodic retrieval are more intimately related than previously assumed.

AUTHOR NOTE

Correspondence concerning this article should be addressed to M. M. Spapé, Institute for Psychological Research and Leiden Institute for Brain and Cognition, Leiden University, Wassenaarseweg 52, Leiden 2333 AK, The Netherlands (e-mail: mspape@fsw.leidenuniv.nl).

REFERENCES

Akçay, Ç., & Hazeltine, E. (2007). Conflict monitoring and feature overlap: Two sources of sequential modulations. *Psychonomic Bulletin & Review*, **14**, 742-748.

Blais, C., Robidoux, S., Risko, E. F., & Besner, D. (2007). Item-specific adaptation and the conflict-monitoring hypothesis: A computational model. *Psychological Review*, **114**, 1076-1086.

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.

Brass, M., Derrfuss, J., & von Cramon, D. Y. (2005). The inhibition of imitative response tendencies: A functional double dissociation of imitative and overlearned responses. *Neuropsychologia*, **43**, 89-98.

Eriksen, B. A., & Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a non-search 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.

Hommel, B. (1998). Event files: Evidence for automatic integration of stimulus–response episodes. *Visual Cognition*, **5**, 183-216.

Hommel, B. (2004). Event files: Feature binding in and across perception and action. *Trends in Cognitive Sciences*, **8**, 494-500.

Hommel, B., Proctor, R. W., & Vu, K.-P. L. (2004). A feature-integration account of sequential effects in the Simon task. *Psychological Research*, **68**, 1-17.

Kerns, J. G., Cohen, J. D., MacDonald, A. W., III, Cho, R. Y., Stenger, A. V., & Carter, C. S. (2004). Anterior cingulate conflict monitoring and adjustments in control. *Science*, **303**, 1023-1026.

Kiesel, A., Kunde, W., & Hoffmann, J. (2006). Evidence for task-specific resolution of response conflict. *Psychonomic Bulletin & Review*, **13**, 800-806.

Kurby, C. A., & Zacks, J. M. (2008). Segmentation in the perception and memory of events. *Trends in Cognitive Sciences*, **12**, 72-79.

Logan, G. D. (1980). Attention and automaticity in Stroop and priming tasks: Theory and data. *Cognitive Psychology*, **12**, 523-553.

Logan, G. D., & Gordon, R. D. (2001). Executive control of visual attention in dual-task situations. *Psychological Review*, **108**, 393-434.

Logan, G. D., & Zbrodoff, N. J. (1979). When it helps to be misled: Facilitative effects of increasing the frequency of conflicting stimuli in a Stroop-like task. *Memory & Cognition*, **7**, 166-174.

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

Mayr, U., & Keele, S. W. (2000). Changing internal constraints on action: The role of backward inhibition. *Journal of Experimental Psychology: General*, **129**, 4-26.

Nieuwenhuis, S., Stins, J. F., Posthuma, D., Polderman, T. J. C., Boomsma, D. I., & De Geus, E. J. (2006). Accounting for sequential trial effects in the flanker task: Conflict adaptation or associative priming? *Memory & Cognition*, **34**, 1260-1272.

Notebaert, W., & Verguts, T. (2008). Cognitive control acts locally. *Cognition*, **106**, 1071-1080.

Simon, J. R., & Rudell, A. P. (1967). Auditory S–R compatibility: The effect of an irrelevant cue on information processing. *Journal of Applied Psychology*, **51**, 300-304.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, **18**, 643-662.

Stürmer, B., Leuthold, H., Soetens, E., Schröter, H., & Sommer, W. (2002). Control over location-based response activation in the Simon task: Behavioral and electrophysiological evidence. *Journal of Experimental Psychology: Human Perception & Performance*, **28**, 1345-1363.

Tzelgov, J., Henik, A., & Berger, J. (1992). Controlling Stroop effects by manipulating expectations for color words. *Memory & Cognition*, **20**, 727-735.

Van Berkum, J. J. A., van den Brink, D., Tesink, C. M. J. Y., Kos, M., & Hagoort, P. (2008). The neural integration of speaker and message. *Journal of Cognitive Neuroscience*, **20**, 580-591.

Waszak, F., Hommel, B., & Allport, A. (2003). Task-switching and long-term priming: Role of episodic stimulus–task bindings in task-shift costs. *Cognitive Psychology*, **46**, 361-413.

Wendt, M., Kluwe, R. H., & Peters, A. (2006). Sequential modulations of interference evoked by processing task-irrelevant stimulus features. *Journal of Experimental Psychology: Human Perception & Performance*, **32**, 644-667.

Wühr, P. (2005). Evidence for gating of direct response activation in the Simon task. *Psychonomic Bulletin & Review*, **12**, 282-288.

Zacks, J. M., & Swallow, K. M. (2007). Event segmentation. *Current Directions in Psychological Science*, **16**, 80-84.

Zacks, J. M., & Tversky, B. (2001). Event structure in perception and conception. *Psychological Bulletin*, **127**, 3-21.

NOTES

1. Note that the same argument holds for manipulations of the frequency of congruent and incongruent trials (e.g., Logan & Zbrodoff, 1979; Tzelgov, Henik, & Berger, 1992)—manipulations that have often been taken to induce adaptive control (Botvinick et al., 1999; Logan, 1980). Making incongruent trials more frequent produces more memory traces for incongruent stimulus–response episodes, which is likely to facilitate responding in such trials for reasons unrelated to control processes.

2. The inclusion of the own-voice condition was motivated by the observation that different neural networks seem to be implicated in tasks inducing conflict per se, such as the Stroop task, and tasks that create conflict between self-generated and externally triggered motor representations, such as when finger movements are signaled by visual displays showing movements of an incongruent finger (Brass, Derrfuss, & von Cramon, 2005). We thought that presenting irrelevant information in the participant's own voice might be similar to the latter condition and were interested to see whether it might increase or otherwise affect conflict-related effects. Given that the voice variable did not yield any significant effect and was not involved in any interaction, we will neglect it in the following study and focus on the repetition and alternation of voice only.

(Manuscript received October 29, 2007;
revision accepted for publication June 11, 2008.)