Journal of Experimental Psychology:
Learning, Memory, and Cognition
2005, Vol. 31, No. 4, 736–749

0278-7393/05/$12.00 DOI: 10.1037/0278-7393.31.4.736

Auditory Attentional Capture During Serial Recall: Violations at Encoding of an Algorithm-Based Neural Model?

Robert W. Hughes
Cardiff University

François Vachon
Université Laval

Dylan M. Jones
Cardiff University

A novel attentional capture effect is reported in which visual–verbal serial recall was disrupted if a single deviation in the interstimulus interval occurred within otherwise regularly presented task-irrelevant spoken items. The degree of disruption was the same whether the temporal deviant was embedded in a sequence made up of a repeating item or a sequence of changing items. Moreover, the effect was evident during the presentation of the to-be-remembered sequence but not during rehearsal just prior to recall, suggesting that the encoding of sequences is particularly susceptible. The results suggest that attentional capture is due to a violation of an algorithm rather than an aggregate-based neural model and further undermine an attentional capture-based account of the classical changing-state irrelevant sound effect.

Keywords: attentional capture, serial recall, irrelevant sound effect, encoding, neural model

The capacity to retain the serial order of incoming information over the short term (i.e., serial memory) has long been recognized as a critical cognitive faculty (e.g., Lashley, 1951). However, even when attention is steadfastly directed toward to-be-remembered (TBR) information, serial memory is susceptible to corruption by the presence of task-irrelevant information, just as is naming the ink color of a color word (Stroop, 1935), reaching for and grasping an object (e.g., Tipper, Howard, & Jackson, 1997), or searching for a visual target (e.g., Yantis, 1998; for a discussion, see Hughes & Jones, 2003b). Because their processing is largely obligatory and the auditory system is acutely sensitive to stimulus change or novelty, irrelevant auditory stimuli are particularly effective at disturbing the efficient execution of a task by exogenously capturing attention[1] (see, e.g., Eimer, Nattkemper, Schröger, & Prinz, 1996). Important advances in the understanding of the attentional processes associated with stimulus change in the auditory modality have been made through the use of psychophysiological indices, such as evoked potentials, that assess the impact of a single stimulus that deviates from a pattern (the *oddball*; for reviews, see Näätänen, 1992; Schröger, 1997). That such effects occur even when the auditory sequence in which the deviation occurs is unattended (the “passive” variant of the oddball paradigm; Eimer et al., 1996) suggests that the flexibility (and lability) of attention can be partly accounted for by the degree of obligatory sensory processing of auditory information.

Relatively little research has been conducted on auditory attentional capture by a single deviant stimulus within the cognitive–behavioral tradition. Although some studies have shown that the ability to respond to target auditory stimuli is impaired when changes occur on a task-irrelevant dimension of those stimuli (e.g., Mondor, Zatorre, & Terrio, 1998; Schröger & Wolff, 1998), in these cases the “irrelevant” changes not only occur within the same modality as the target stimuli, they also occur within the same stimulus as that which acts as the target (e.g., Mondor et al., 1998; Schröger & Wolff, 1998) or at least within the same (attended) auditory sequence that contains the target (Dalton & Lavie, 2004). As such, these studies do not speak to whether irrelevant changes

[1] It may be worth making explicit a distinction between two ways in which the term *capture* has been used in the auditory perception–attention literature. An incoming stimulus that is integrated into an unfolding perceptual stream by virtue of its similarity in timing or composition to the other stimuli in that stream may be described as being perceptually captured by that stream (e.g., Bregman, 1990). For example, Nicholls and Jones (2002) were able to use a sequence with regularly presented irrelevant stimuli that were similar in timing and pitch to a redundant end-of-list suffix in order to capture the suffix into the irrelevant stream so that the suffix was no longer grouped with the to-be-remembered list (and, thereby, its impact was much reduced). *Attentional capture*, on the other hand, refers to a stimulus that alters the direction of attention away from the prevailing focus. Indeed, the suggestion we make in the present article is that attentional capture operates in the opposite fashion to perceptual capture: It is an incoming stimulus that does not conform to an algorithm governing the integration of stimuli into an extant stream (i.e., it is not perceptually captured by the stream) that assumes the power to capture attention.

Robert W. Hughes and Dylan M. Jones, School of Psychology, Cardiff University, Cardiff, United Kingdom; François Vachon, École de Psychologie, Université Laval, Québec City, Québec, Canada.

The research reported in this article received financial support from Grant RES-000-22-0959 from the United Kingdom’s Economic and Social Research Council awarded to Dylan M. Jones and Robert W. Hughes. François Vachon worked on the study during a postgraduate internship, for which he received financial support from the Fonds Québécois pour la Recherche sur la Nature et les Technologies and the Natural Sciences and Engineering Research Council of Canada.

Correspondence concerning this article should be addressed to Robert Hughes, School of Psychology, Cardiff University, Cardiff CF10 3AT, United Kingdom. E-mail: hughesrw@cardiff.ac.uk

that occur within a stream of auditory information that is never-to-be-attended can capture attention. In the present study, we sought to (a) examine whether attentional capture by a single deviant stimulus—embedded in a sequence that is, nominally, never-to-be-attended—can be demonstrated within the context of a simple, visually presented, short-term serial memory task; (b) tease apart which of two broad approaches to attentional capture best accounts for any observed capture effects; and (c) establish what features of the serial memory task are susceptible to disruption.

An Aggregate-Model Approach to Attentional Capture

We begin with the uncontroversial assumption that attentional capture by a stimulus is some product of the degree to which its features depart from immediate prior experience or expectations, what may broadly be called a *neural model*. This idea is embodied in classic (e.g., Sokolov, 1963) and contemporary (e.g., Cowan, 1995; Schröger, 1997) accounts of the so-called *orienting response* (OR), a complex of behavioral responses (e.g., slowing of the heart rate, increase in skin conductance) that may (but need not necessarily) accompany the covert phenomenon of attentional capture.

What commands less agreement is the nature of the neural model. One view is that the model is an aggregation of the stimulus sequence, a prototypical representation of the stimuli making up the sequence. Attentional capture is some product of the mismatch between a given stimulus and the aggregate model of its predecessors. Thus, a sequence in which the same item is repeated, for example *4 4 4 4* . . . , would quickly give rise to a definitive depiction and each *4* would, correspondingly, be progressively less likely to capture attention (i.e., the OR habituates; e.g., Cowan, 1995; Sokolov, 1963). In contrast, in an auditory sequence in which each item is different, for example, *4 1 3 7* . . . , each item should capture attention because each would mismatch the aggregate model of its predecessor(s).

Not only is this aggregate account of attentional capture attractive by virtue of its simplicity, it is attractive also because it may provide an account of auditory distraction effects that have not typically been ascribed to an attentional capture mechanism. For example, it may help to understand the so-called *changing-state irrelevant sound effect* (Colle & Welsh, 1976; D. M. Jones & Macken, 1993; D. M. Jones, Madden, & Miles, 1992; Salamé & Baddeley, 1982, 1989; for an overview, see Hughes & Jones, 2001). In this paradigm, participants are asked to try to recall a list of visual–verbal items in serial order while being exposed to background sound that they are told to ignore. Despite the fact that the serial recall task is presented in the visual modality, deficits of up to 20%–50% can be found with some types of irrelevant sound (e.g., narrative speech; see Ellermeier & Zimmer, 1997, for a discussion of effect size). The empirical signature of this irrelevant sound effect is the *changing-state effect*, whereby a sequence with changing components invariably produces more disruption than one with repeated components (e.g., Jones et al., 1992).

The aggregate-model account of attentional capture would seem to provide a compelling and parsimonious explanation of the changing-state irrelevant sound effect. On this account, as noted, each stimulus in a changing-state (CS) sequence (e.g., *4 1 3 7* . . . .) would have the power to capture attention, and such repeated capture is likely to disrupt the prevailing serial recall task. In contrast, in a steady-state (SS) sequence (e.g., *4 4 4 4*. . .), the power of each stimulus to capture attention would soon habituate (Cowan, 1995; Elliott, 2002) thereby sparing serial recall performance.

Two interesting predictions—with implications for understanding the character of attentional capture generally—flow from this way of interpreting the CS effect. First, and significantly, it predicts that a single deviant stimulus within the irrelevant sequence will materially depress serial recall performance, a prediction that has, somewhat surprisingly, never been tested in the irrelevant sound paradigm. Second, it predicts that a deviant stimulus—for example, a single item presented out of rhythm with the others—will have a greater attention-capturing effect when it is part of a sequence with a repeated item than when it is part of a sequence with changing items.[2]

According to the aggregate-model approach, each item in a CS sequence, for example *4 1 3 7 2 6 8 5,* has the power to capture attention because each item deviates from the item-based prototype of its predecessors. This means that the introduction of a temporal deviation into this sequence—for example, delaying the fifth item, that is *4 1 3 7_ **6**8 2 5*, could not capture attention any further. In stark contrast, each item in the sequence *4 4 4 4 4 4 4 4* has less and less power to capture attention. Given that the fifth item in this regular SS sequence would be unlikely to capture attention, delaying that item, that is *4 4 4 4_**4**4 4 4*, should indeed have such power. In other words, within the context of a CS sequence, delaying the fifth item cannot produce an attentional capture effect because that item—by virtue of its item-based deviation from its predecessor(s)—captures attention regardless of whether or not it is delayed. Within the context of an SS sequence, however, delaying the fifth item can and should produce an attentional capture effect because that item does not already have attention-capturing power.

The foregoing analysis relies on the possibly simplifying assumption that attentional capture is an “all-or-nothing” phenomenon (i.e., the attentional capture mechanism is either elicited or it is not; it cannot be partially elicited). However, even if we assume that attentional capture can be elicited to different degrees (but where there is an upper limit), the aggregate model still predicts some difference in the degree to which a temporal deviant will have an impact according to whether it is embedded in an SS or CS sequence. This is because, in a CS sequence, delaying the fifth item can only exacerbate the degree to which attentional capture is already elicited by that item (by virtue of its item-based deviation). Within an SS sequence, in contrast, the temporal deviant would alter the situation from one in which the attentional capture mechanism is not elicited to one in which it is. It seems reasonable to suggest that the triggering of the attentional capture mechanism from a baseline level should be more disruptive than merely increasing the degree to which that mechanism is engaged.

[2] Note that, in the present article, the terms *changing state* and *steady state* are used exclusively to refer, respectively, to a sequence containing a succession of items changing in acoustic composition from one to the next and a sequence made up of a single repeating item and will not be used to refer to the timing of the items (for which the terms *irregular* and *regular* will be used).

An Algorithm-Model Approach to Attentional Capture

On a different approach to attentional capture, the second of the foregoing predictions does not hold. Instead of construing the model (on which capture is based) as an aggregate, we may think of it as an algorithm (a view more in line with later formulations of the OR, e.g., Sokolov, 1975). On this account, attentional capture results from a violation of a pattern, or rule, that governs the organization of the stimulus sequence. An algorithm-model account provides an explanation for several examples of attentional capture that are problematic for an aggregate-model account. For example, a violation in canonical order such as *1 2 3 4 5 7 8 9,* and deviations from locally established rules (e.g., *5 4 3 8 7 7 4 6*) produce appreciable capture effects as indexed by physiological indices of the OR (Unger, 1964; see also Velden, 1978). In neither of these cases would the *7* (the second instance of *7* in the latter sequence) have mismatched an aggregate model of the stimuli preceding it any more than any other stimulus in the sequence; it would have violated only a rule that transcends the nature of the individual stimuli themselves. Similarly, results from studies that have used event-related potential (ERP) measures suggest that a psychophysiological response normally associated with attentional capture (the mismatch negativity [MMN] response) occurs in response to deviations based on second-order (based on conjunction) and third-order (based on expectation) characteristics of a sequence (for an overview, see Schröger, 1997). For example, the following types of deviation elicit the MMN response: a pair of tones descending in frequency following a sequence of ascending tones (Saarinen, Paavilainen, Schröger, Tervaniemi, & Näätänen, 1992), repetitions within a sequence of otherwise alternating low and high tones (Nordby, Roth, & Pfefferbaum, 1988), and a pair of same-frequency tones exhibiting a deviant spatiotemporal order (Schröger & Wolff, 1996). Again, such complex deviation effects appear to occur irrespective of whether or not the sequence is being attended (Eimer et al., 1996; but see Woods, 1990).

An algorithm-model approach to attentional capture makes a different prediction from the aggregate-model account with regard to the possible impact on serial recall of a temporal irregularity within an SS compared with a CS irrelevant sequence. First, note that according to the algorithm-model account, in both a standard SS and standard CS sequence (i.e., in which the timing of the items is regular), no item violates a neural model; each item in a CS sequence fits the algorithm of "each item differs from the previous one" and each item in an SS sequence fits the algorithm "each item is the same." On this account, therefore, the classical CS effect cannot be ascribed to an increased propensity for capture in a CS compared with an SS sequence (see General Discussion for an alternative, noncapture-based, explanation of this effect). Second, on this account, if a deviation in timing is introduced, such a deviation should have roughly the same impact on short-term serial recall whether it occurs within an SS or CS sequence. This is because in both cases the item that is out of time with its predecessors is the only one in the sequence to violate any algorithm encapsulated in the model (i.e., that pertaining to the timing of the items) and should therefore be the first and only event to have the power to capture attention.

The central purpose of the experiments reported here, then, is to help characterize the nature of attentional capture. By studying the impact of a temporally deviant stimulus in the context of either repeated or changing stimuli, we hope to help adjudicate which class of explanation—the aggregate or algorithmic type—is most appropriate for understanding attentional capture generally. At the same time, the experiments should inform our understanding of the irrelevant sound effect and particularly the extent to which we may ascribe the CS effect to repeated attentional capture.

Experiment 1

Experiment 1 is, to our knowledge, the first to attempt to establish whether a single item presented irregularly in an otherwise regular sequence of irrelevant sound items has the capacity to disrupt serial recall. Because the sequence itself could contain tokens (words) that were either repeated or were changing, two further predictions can be made: First, the classical CS irrelevant sound effect should be replicated (e.g., Jones et al., 1992). That is, ignoring for the moment those conditions in which a deviation in timing is introduced into a sequence, a CS sequence should produce more disruption than an SS sequence. The second, novel prediction pertains to the veracity of the aggregate-model account of attentional capture and its applicability to the classical CS effect (e.g., Cowan, 1995). If the impact of the deviant stimulus is greater when embedded in a sequence of repeated sounds than in a sequence of changing sounds, then aggregation type models would be strengthened appreciably and an attentional capture account of the classical CS effect would remain viable (but see General Discussion). However, if the impact of the deviant is roughly the same, regardless of irrelevant sequence context, then algorithmic approaches to attentional capture should be favored and an account of the classical CS effect couched in terms of attentional capture seems implausible.

Method

Participants. Twenty-six psychology students at Cardiff University took part in the experiment for either course credit or a small honorarium. All reported normal or corrected-to-normal vision and hearing.

Apparatus and materials. The TBR visually presented sequences were eight items length and were taken without replacement from the digit-set *1–9* and arranged in a quasi-random order with the constraint that successive digits were not adjacent integers. Each item was presented sequentially in a 72-point Times New Roman font at the center of the screen of a Macintosh Performa with *PsyScope 1.2.2* software (Cohen, MacWhinney, Flatt, & Provost, 1993). Each digit was presented for 350 ms and the interstimulus interval (ISI; offset to onset) was 400 ms.

For the irrelevant auditory sequences, the spoken digits *1–9* were recorded in a female voice at an approximately even pitch, and each was then digitally edited with the *SoundForge 7* software (Sony Media Software, 2004) to last 250 ms. These were presented stereophonically at approximately 65 dB(A). Using these stimuli, we generated the following four types of irrelevant sequences (see Figure 1A):

1. An SS sequence consisted of one of the auditory digits repeated eight times with a regular ISI of 500 ms. The digit was chosen randomly for each trial, with the constraint that it was always one that was also in the accompanying TBR list. The onset of each auditory digit preceded that of each of the eight visual digits by 75 ms to produce approximate phenomenal simultaneity (this was the case for each auditory stimulus in each condition except for any ISI deviant stimulus; see Sequence Types 3 and 4 below).

2. A CS sequence consisted of eight auditory digits (the same eight

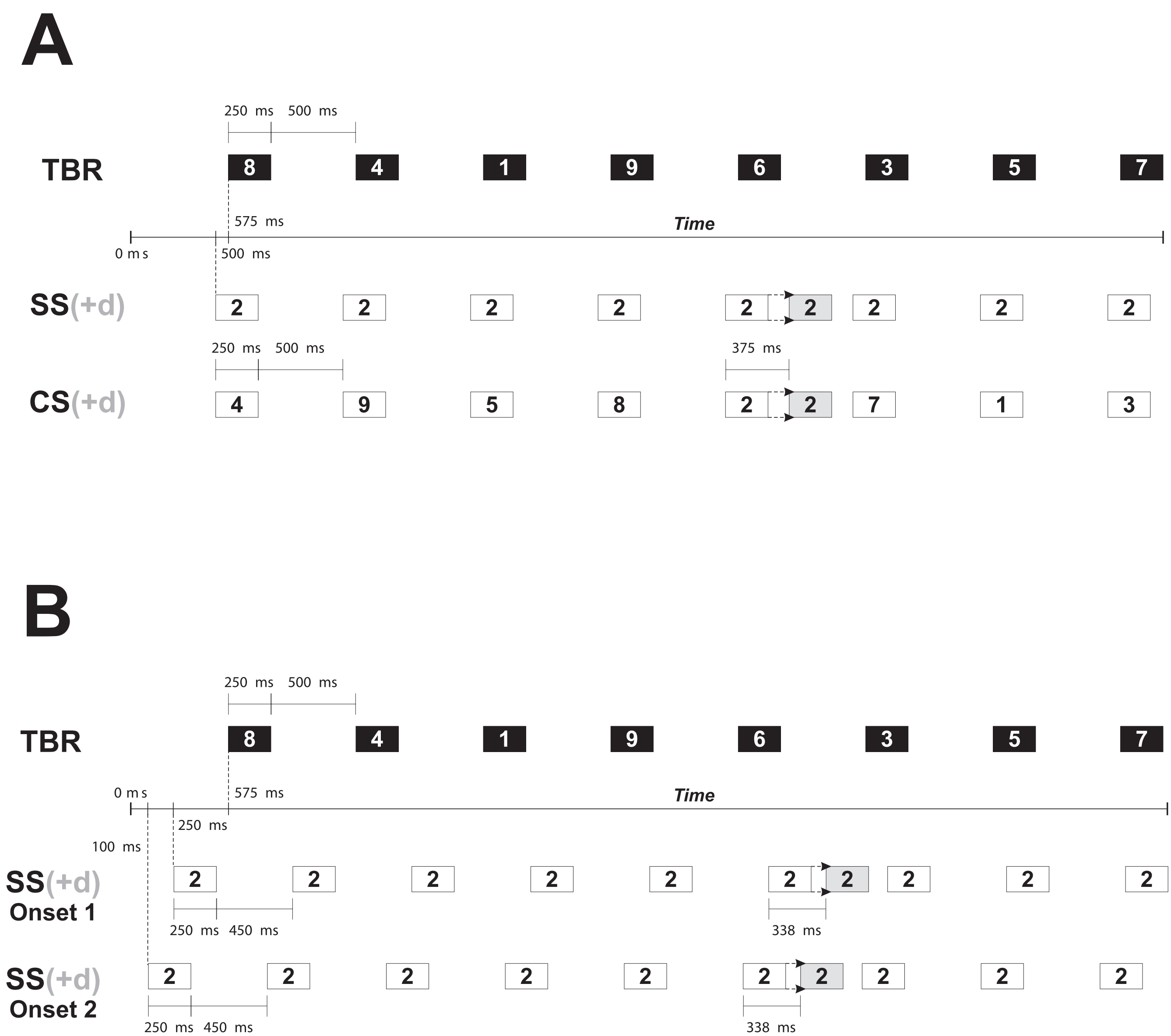

Figure 1. A: A schematic representation of stimulus presentation in Experiments 1, 3, and 4 (note that the to-be-remembered [TBR] items were letters rather than digits in Experiment 4). B: Timings used for Experiment 3. SS = steady state; CS = changing state. The SS + deviation (*d*) and CS + *d* conditions are represented within the representations of the SS and CS sequences as a shift in the temporal position of the fifth irrelevant item.

digits that made up the accompanying TBR list but presented in a different order), again presented with a regular ISI of 500 ms.

3. An SS + ISI deviation (*d*) sequence was identical to an SS sequence, except the presentation of the fifth auditory item was delayed by 375 ms. That is, the ISI between the fourth and fifth items was 875 ms, and that between the fifth and sixth was therefore 125 ms.

4. A CS + ISI *d* was identical to a CS sequence, except it included the same ISI deviation as described for the SS + ISI *d* sequence.

In both CS conditions, we took care to ensure that a concurrent auditory and visual digit were not the same digit and that there were no obvious ascending or descending runs of three or more digits either within the auditory sequence or across the auditory and visual sequences.

Design. A repeated-measures design was used with two factors: serial position (eight levels) and auditory condition (five levels, corresponding to the four sequence types described above with a quiet control condition). The auditory conditions were presented pseudorandomly, with the constraint that there were no immediate repeats of the same condition and each condition was presented once every five trials.

Procedure. Participants were tested individually in a sound-attenuated booth. Each participant read standard instructions that informed them of the nature of the serial recall task involved and asked them to ignore any sounds they may hear over the headphones. Participants were also informed that the trials would be presented at a preset pace: 50 ms following the offset of the last visual item the screen flashed from white to black for 150 ms, which signaled to the participants that they should begin to write out the TBR list. From the offset of the screen flashing, there were 15 s before the presentation of the first item of the next TBR list. A 500-ms tone was presented over the headphones 13 s into the 15 s of "writing time" to signal the participants that the presentation of the first item of the next sequence was imminent. The experimental block consisted of 15 trials for each of the five auditory conditions making 75 trials in all. Two practice trials (1 quiet trial and 1 SS trial) were given prior to the experimental block proper. The experiment lasted 35 min.

Results

The raw serial recall data were scored according to a strict serial recall criterion: Recalled items were only scored as correct if they corresponded to their presentation position (as in all experiments reported in this article). Figure 2 shows the mean percentage of items correctly recalled in each of the five auditory conditions across the eight serial positions. The figure shows compellingly that performance is generally depressed in the CS conditions compared with the SS conditions, regardless of the ISI deviation manipulation and also that performance is poorer in the SS + ISI *d* (M = 74.01, SE = 3.10) than in the SS condition (M = 76.89, SE = 3.01) and poorer in the CS + ISI *d* (M = 60.80, SE = 3.50) than in the CS condition (M = 65.80, SE = 3.02).

An initial repeated-measures analysis of variance (ANOVA) incorporating all five auditory conditions revealed a main effect of auditory condition, $F(4, 100) = 26.20$, $MSE = 10.32$, $p < .001$, with planned contrasts showing that all irrelevant speech conditions except the SS condition produced poorer performance than did the quiet control (all ps $< .05$). Given that interest centered on the four conditions in which irrelevant sound was presented (i.e., SS, SS + ISI *d*, CS and CS + ISI *d*), the data from these four conditions were then analyzed with a 2 (state: steady or changing) × 2 (deviation: present or absent) × 8 (serial position) ANOVA (i.e., the data from the quiet condition were omitted). This analysis revealed a main effect of state, $F(1, 25) = 53.14$, $MSE = 13.87$, $p < .001$, a main effect of deviation, $F(1, 25) = 7.53$, $MSE = 7.79$, $p < .01$, and a main effect of serial position, $F(7, 175) = 28.29$, $MSE = 9.15$, $p < .001$. There were no significant interactions between any of the three factors; most notably, there was no indication of a State × Deviation interaction, $F(1, 25) = 0.429$, $MSE = 4.71$, $p = .52$ (Cohen's $d = -0.262$; power to detect a medium effect size [0.50] = .23; note that the negative d value reflects the fact that, although the experiment was relatively low in power, the nonsignificant trend—in contradiction to the prediction of the aggregate model account—is for a greater deviation effect in the CS context}. Thus, the deviant had roughly the same degree of impact regardless of whether it was embedded in an SS or CS irrelevant sequence.

Discussion

The results of Experiment 1 are relatively straightforward: As well as replicating the classical CS effect, they demonstrate for the first time that an irrelevant sequence containing a single deviation in the ISI, regardless of whether that sequence is made up of CS or SS stimuli, incurs an additional cost to serial recall. Moreover, as reflected in the absence of an interaction between deviation and serial position, a temporal deviant appears to exert a general effect on the serial recall of the TBR list; the disruption is not localized to the TBR items closest in time to the deviation (items 5 and 6). Of course, this is not surprising; a disturbance to the encoding and/or rehearsing of the list as it is presented is likely to cause a propagation and back propagation of errors throughout the list. Moreover, different participants, and the same participants at different times, may be at different points in their rehearsal cycles when the temporal deviant occurs, which would also serve to result in a diffusion of the effect across the curve (see also the *Discussion* of Experiment 3).

We would argue that the deviation effect is the result of attention being captured from the primary task. However, given that the effect of the deviant stimulus was roughly the same in a repeated and changing context, the idea that the phenomenon of attentional capture is the product of a departure from some aggregate representation of the stimulus sequence seems implausible. On this account, each irrelevant item in the CS condition would already be

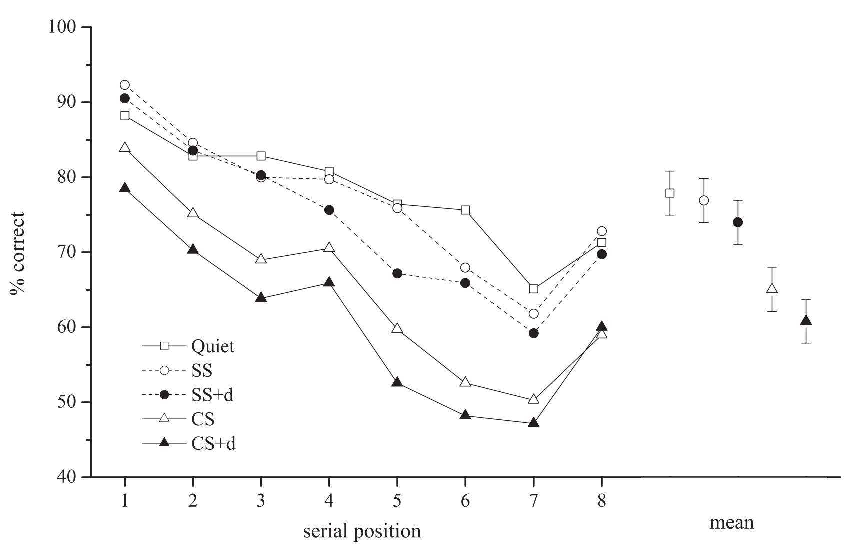

Figure 2. Mean percentage of items correctly recalled in the quiet, steady-state (SS), steady-state + interstimulus interval (ISI) deviation (SS + ISI *d*), changing-state (CS), and changing-state + ISI deviation (CS + ISI *d*) conditions in Experiment 1. Error bars represent 95% confidence intervals.

capturing attention, and hence the additional deviation in timing should not have had much, if any, additional impact in this condition. In turn, the finding casts doubt on an attentional capture account of the classical CS effect. In contrast, the results are compatible with an algorithm-model account of attentional capture. The deviation in timing would have violated the algorithm of "each item arrives about half a second after the previous one" and hence would have the propensity to capture attention. However, there was no reason to have expected the effect to be diminished in the CS condition because, on this account, each CS item would not have been already endowed with attention-capturing power.

It seems possible that the action of a single deviant stimulus is confined to the encoding (or perhaps more properly, the assembly) of the rehearsal cohort. This leaves open the possibility that once the rehearsal cohort is established the deviant stimulus would no longer be disruptive. This mode of action would contrast sharply with the classical CS irrelevant sound effect, which is as marked typically when the irrelevant sound is presented following the presentation of the TBR items (i.e., during a retention interval, at which time participants are expected to keep rehearsing the TBR sequence) as it is when the sound is presented during the presentation of the TBR items (Macken, Mosdell, & Jones, 1999; Miles, Jones, & Madden, 1991). In Experiment 2, therefore, we examined whether presenting the deviation within an irrelevant sequence presented only during a retention interval following the presentation of the final TBR item would produce a deviation effect like that observed in Experiment 1.

Experiment 2

If the deviation effect observed in Experiment 1 was indeed reliant on the concurrent presentation of the irrelevant and TBR sequences, then in Experiment 2, performance should not be depressed in the SS + ISI *d* condition compared with the SS condition nor depressed in the CS + ISI *d* condition compared with the CS condition. However, in line with previous findings (Macken et al., 1999; Miles et al., 1991), we should indeed observe a CS effect; that is, performance should be poorer in the CS than in the SS condition.

Method

Participants. Twenty-nine psychology students at Cardiff University took part in the experiment for either course credit or a small honorarium. All reported normal or corrected-to-normal vision and hearing.

Apparatus and materials. These were identical to those used in Experiment 1, except for the following details: The signal for the participant to begin writing out the TBR sequence (the screen flashing) was delayed for 7 s. In the irrelevant sound conditions, the first auditory digit was presented 400 ms following the last TBR item and the offset of the final auditory digit occurred 50 ms before the cue to recall the TBR list (the screen flashing).

Design. The design was identical to that used in Experiment 1.

Procedure. This was the same as in Experiment 1, except the participants were told that following the final TBR digit, they should keep covertly rehearsing the sequence until the screen flashed some 7 s later and that they should ignore any sounds they might hear during this retention interval. The experiment lasted 55 min.

Results

Figure 3 shows the mean percentage of items correctly recalled in the five conditions across the eight serial positions in Experiment 2. First, it is evident that performance is depressed in the CS conditions (with or without a deviation) compared with SS conditions (with or without a deviation), although this CS effect is not as compelling as that obtained in Experiment 1. Second, and most important, in contrast to Experiment 1, there is no evidence that a temporal deviation within either an SS sequence (SS + ISI *d*: M = 76.35, SE = 2.96, compared with SS: M = 75.67, SE = 3.26) or a CS sequence (CS + ISI *d*: M = 72.67, SE = 3.31, compared with CS: M = 71.67, SE = 3.13) incurred an additional cost to serial recall. An initial repeated-measures ANOVA incorporating all five conditions in Experiment 2 revealed a main effect of auditory condition, $F(4, 112) = 12.45$, $MSE = 5.37$, $p < .001$, and planned

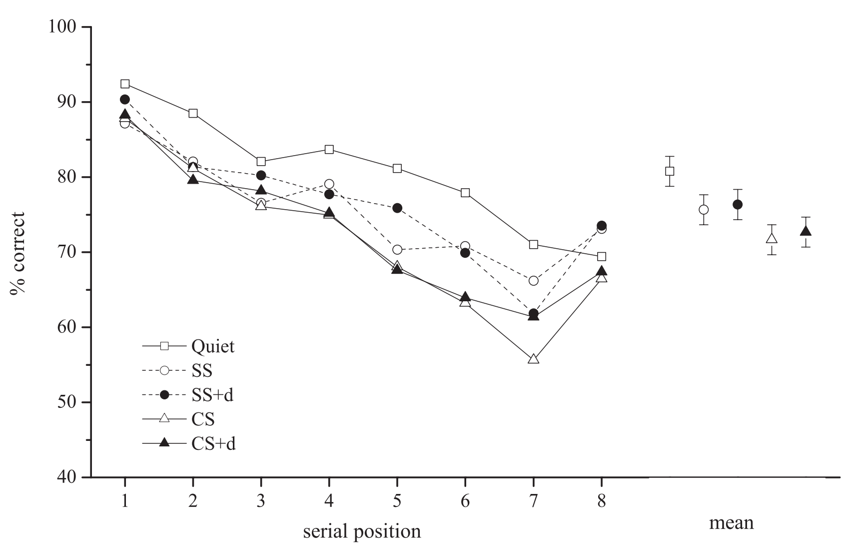

Figure 3. Mean percentage of items correctly recalled in the quiet, steady-state (SS), steady-state + interstimulus interval (ISI) deviation (SS + ISI *d*), changing-state (CS), and changing-state + ISI deviation (CS + ISI *d*) conditions in Experiment 2. Error bars represent 95% confidence intervals.

contrasts showed that performance was reliably poorer in all irrelevant speech conditions compared with the quiet condition (all $ps < .05$).

The data were then submitted to a 2 (state) × 2 (deviation) × 8 (serial position) repeated measures ANOVA, as in Experiment 1 (with the data from the quiet condition omitted). This analysis revealed a main effect of state, $F(1, 28) = 16.37$, $MSE = 5.69$, $p < .001$, a main effect of serial position, $F(7, 196) = 13.78$, $MSE = 15.01$, $p < .001$, and a significant State × Serial Position interaction, $F(7, 196) = 2.97$, $MSE = 1.36$, $p < .01$, which may reflect the fact that the CS effect was weaker at primacy. That the CS effect is weaker at primacy has been observed in previous studies (e.g., D. M. Jones & Macken, 1993, 1995), but no theoretical significance is attached to this interaction because it more than likely reflects a scalar effect. What is more important, as suspected on the basis of Figure 3, the main effect of deviation was not significant, $F(1, 28) = 1.35$, $MSE = 2.77$, $p = .25$ (Cohen's $d = -0.44$); again, the power of this experiment to detect a medium size effect was quite low (.26), however, as indicated by the negative d value, the nonsignificant trend was for better performance in the deviant conditions. Moreover, there was no interaction between deviation and state, $F < 1$. A three-way State × Deviation × Serial Position interaction emerged, $F(1, 28) = 2.11$, $MSE = 1.69$, $p < .05$. Again, we do not attempt to ascribe any functional significance to this interaction.

Discussion

In Experiment 2, in which the irrelevant sequence was presented only during a retention interval as opposed to during the presentation of the TBR items as in Experiment 1, the usual CS effect was evident thus replicating a key finding in the irrelevant sound literature (D. M. Jones & Macken, 1993; Miles et al., 1991). However, the CS effect was less compelling than in Experiment 1, the reason for which is unclear. One previous study found no difference in the impact of irrelevant sound presented during presentation compared with during a retention interval (Miles et al., 1991), whereas another found evidence for somewhat greater disruption when the sound was presented during the retention interval (Macken et al., 1999). Given these mixed results, the issue of whether and how the magnitude of the classical irrelevant sound effect is modulated by the point at which it is presented—although tangential to the goals of the present study—would seem to warrant further research.

The key result of Experiment 2, as far as our present goals are concerned, is that although there was a reliable CS effect, there was no evidence of a deviation effect. It may be the case, then, that a temporally deviant stimulus selectively disrupts the process of assembling or assimilating the items into an initial rehearsal cohort during the presentation of TBR material; there is no evidence that it disrupts the process of continuing to rehearse the TBR items during a retention interval. This contrast suggests that the classical CS effect and the present deviation effect may be functionally distinct (see General Discussion).

One suggestion as to why the deviation effect only has an effect during presentation could be based on the fact that, in our method, there was a very obvious or highly transparent temporal relationship between the irrelevant and relevant items (up to the point at which the deviant occurred). Let us assume, first, that the regularity within the TBR list aids the process of assimilating each incoming TBR item into an initial rehearsal cohort. This might, for example, be because selective attention can be pulsed according to that rhythm (i.e., each TBR item is more efficiently selected as a result of the economical pulsing of attention during times at which items are expected to occur; cf. M. R. Jones, 1976, 2001) or because it aids in grouping the TBR items into a coherent stream (see, e.g., Nicholls & Jones, 2002). If so, having a high degree of between-sequence temporal transparency may facilitate the process of selecting the TBR items over the irrelevant items. If this is the case, a temporal deviation in the irrelevant sequence may corrupt the assimilation process. In contrast, a deviation in the irrelevant sequence may not disturb rehearsal during the retention interval because continuing to rehearse the TBR items would not be slaved to the rhythm of any external events. In other words, the impact of the deviant may stem not so much from its violation of a neural model of the irrelevant sequence per se, but from some neural model that embodies the temporal relationship between the relevant and irrelevant stimuli. In Experiment 3, therefore, we explored whether the deviation effect is attenuated, if not abolished, if the temporal transparency between the relevant and irrelevant sequences is sharply reduced.

Experiment 3

In this experiment, we reduced the degree to which the temporal relationship between the irrelevant and relevant sequence was transparent to the participants in two ways. First, the ISI in the irrelevant sequence was shortened from 500 ms to 450 ms so as to remove the systematic synchronicity between the relevant and irrelevant items (note that this also increased the number of irrelevant items from eight to nine). Second, the onset time of the irrelevant sequence was varied across trials (with two levels; see Figure 1B and the following *Method* section for details). If the deviation effect observed in Experiment 1 was reliant on the high temporal transparency in the timing of the irrelevant and relevant sequences then, in Experiment 3, the effect should be attenuated markedly, if not abolished.

It is important to note that the size of the deviation (when present) was equivalent proportionally to that used in Experiments 1 and 2. That is, although the delay between irrelevant items 4 and 5 was, in absolute terms, slightly reduced compared with that in Experiments 1 and 2, the size of the deviation relative to the ISI between irrelevant items 1–4 remained the same. As can be seen in Figure 1B, another inevitable consequence of the changes made for this experiment was that when a deviant occurred it did so at a slightly later point relative to the TBR list compared with Experiments 1 and 2 and could occur at two slightly different points relative to the TBR list. Note, however, that the deviant still occurred at broadly the same point regardless of the changes; the deviant always occurred between TBR items 5 and 6.

Given that the goal of Experiment 3 was to ascertain whether between-sequence temporal transparency played a role in the deviation effect found in Experiment 1, we simply contrasted performance under a SS and SS + ISI d condition and a quiet control condition (we sought to replicate the deviation effect in both SS and CS contexts in Experiment 4).

Method

Participants. Twenty-eight Cardiff University students reporting normal or normal-to-corrected vision and normal hearing took part in return for course credits. None of the participants had taken part in Experiments 1 or 2.

Apparatus and materials. The apparatus and materials were the same as in Experiments 1 and 2, except for the following details: For the SS sequence, a randomly chosen auditory digit (again with the constraint that the digit was always one present in the accompanying TBR list) was repeated nine times with a regular ISI of 450 ms, yielding a presentation rate slightly faster than in the TBR visual sequence. The SS + ISI *d* sequence was identical to an SS sequence except that the fifth auditory digit was delayed by 338 ms. In such a sequence, the ISI between the fourth and fifth digits was 788 ms and that between the fifth and the sixth digits was 112 ms.

Design. A 2 (auditory sequence onset) × 3 (auditory condition) × 8 (serial position) repeated-measures design was used. The three auditory conditions correspond to the two types of sequence described above (SS and SS + ISI *d*) and a quiet control condition. The auditory sequence could have two start points (or onsets) in order to reduce the predictability of the timing between the relevant and irrelevant items. Specifically, the onset of the first auditory item could precede the onset of the first TBR item by either 250 ms (Onset 1) or 100 ms (Onset 2). Clearly, the onset factor could not be manipulated in the quiet condition and so a full factorial design was not possible. The five conditions (quiet, $SS_{Onset\ 1}$, $SS_{Onset\ 2}$, SS + ISI $d_{Onset\ 1,}$ and SS + ISI $d_{Onset\ 2}$) were presented pseudorandomly within a block of trials with the constraint that each condition was presented once every six trials, except for the quiet condition which was presented twice every six trials.

Procedure. The procedure was identical to that used in Experiments 1 and 2, except for the following details: The experiment consisted of 60 experimental trials preceded by two practice trials (one quiet trial and one SS trial). There were 10 trials in each of the four conditions involving an irrelevant sound sequence ($SS_{Onset\ 1}$, $SS_{Onset\ 2}$, SS + ISI $d_{Onset\ 1,}$ and SS + ISI $d_{Onset\ 2}$) and 20 trials in the quiet condition. The experiment lasted approximately 25 min.

Results

Figure 4 shows the mean number of items correctly recalled in the quiet, SS, and SS + ISI *d* conditions (the data for the latter two conditions have been collapsed across the two levels of the Onset factor because this factor did not influence the results, as described below). There is clear evidence that the deviation effect was replicated despite the marked reduction in the degree of between-sequence temporal transparency in this experiment (SS + ISI *d*: $M = 71.21$, $SE = 3.29$, compared with SS: $M = 73.93$, $SE = 3.01$). We first subjected the data to a 2 (onset) × 2 (deviation; SS vs. SS + ISI *d*) × 8 (serial position) repeated-measures ANOVA in order to examine whether the impact of the deviant differed according to the two levels of the onset factor. There was no main effect of onset, $F(1, 27) = 3.82$, $MSE = 3.41$, $p > .05$, but there was a main effect of deviation, $F(1, 27) = 5.52$, $MSE = 3.01$, $p < .05$, and a main effect of serial position, $F(7, 189) = 13.69$, $MSE = 7.72$, $p < .01$. There were no significant interactions between any of the three factors; most notably, there was no Deviant × Onset interaction, $F(1, 27) = 1.24$, $MSE = 2.82$, $p > .05$.

For the main analysis, therefore, the data for the SS and SS + ISI *d* conditions were collapsed across the Onset factor. A 3 (auditory condition: quiet, SS and SS + ISI *d*) × 8 (serial position) repeated-measures ANOVA revealed a main effect of auditory condition, $F(2, 54) = 3.70$, $MSE = 12.39$, $p < .05$, and a main effect of serial position, $F(7, 189) = 13.47$, $MSE = 20.04$, $p < .01$. The interaction between auditory condition and serial position also just reached significance, $F(14, 378) = 1.72$, $MSE = 2.41$, $p = .05$. This interaction may have arisen in large part as a result of the peak evident at Serial Position 4 in the quiet condition. Corroborating this, the interaction disappeared when the quiet condition was removed from the analysis, $F < 1$. Planned contrasts revealed further that performance in the quiet and SS conditions did not differ reliably, $F(1, 27) = 1.36$, $MSE = 20.53$, $p > .05$, but that performance was indeed reliably poorer in the SS + ISI *d* condition compared with the SS condition, $F(1, 27) = 5.52$, $MSE = 12.01$, $p < .05$.

Discussion

In Experiment 3, the deviation effect was replicated and did not appear in an attenuated form despite a marked reduction in the degree of between-sequence temporal transparency compared to

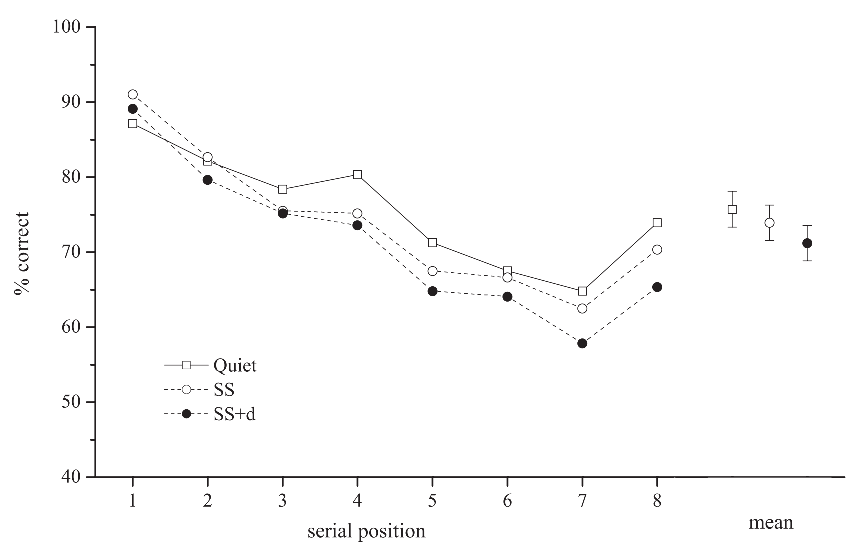

Figure 4. Mean percentage of items correctly recalled in the quiet, steady-state (SS), and steady-state + interstimulus interval (ISI) deviation (SS + ISI *d*) in Experiment 3. Note that the data from the SS and SS + ISI *d* conditions are collapsed across the Onset factor. Error bars represent 95% confidence intervals.

that present in Experiment 1. The main motivation for Experiment 3 was the observation that a deviation effect was not observed when the deviant was presented during a postencoding (or post-assimilation) retention interval (Experiment 2). It was speculated that the disruptive power of the deviant in Experiment 1 may have resided in its power to disturb a TBR item-assimilation process that was aided by the otherwise synchronous and highly transparent temporal relationship between the relevant and irrelevant items. The present result suggests that this may not have been the case. However, it seems prudent to be cautious about this conclusion without further converging evidence: Participants may have become sensitive, at least at an implicit level (e.g., Seger, 1994), to the temporal contingency that still existed between the relevant and irrelevant sequences. Such sensitivity may at least have contributed to the deviation effect found in Experiment 3. To address this issue fully, however, would require the focus of a future study that systematically and parametrically varied the degree of between-sequence temporal transparency.

The result of Experiment 3 does allow us to conclude that the deviation effect does not seem to be reliant on a highly obvious temporal relationship between the relevant and irrelevant items. This in turn suggests that our speculative account of why the effect did not obtain when presented during a retention interval (Experiment 2) might be incorrect. An alternative possibility is that it is even more likely during the retention phase than during the presentation phase that different participants (and perhaps the same participants on different trials) may be at different points in their rehearsal cycles when the deviant occurs. Assuming that some portions of the rehearsal cohort are more vulnerable to disruption by an irrelevant temporal deviant than others, the effect may have been diluted to a point at which it was not detectable in the data. Perhaps a future study in which participants' (whispered) rehearsal activity is recorded may shed some light on this possibility. For the present, however, we step back from our interest in the possible role of between-sequence temporal transparency and turn to the final experiment to try to replicate a key outcome of Experiment 1, namely, that the temporal deviant had roughly the same impact within a CS sequence as in an SS sequence. Moreover, in Experiment 4, we sought to extend the generality of the deviation effect by examining whether it obtains when the irrelevant and relevant items are taken from different verbal categories.

Experiment 4

Given that the issue of attentional capture by a single deviant had not previously been addressed within the irrelevant sound setting, it was a rather arbitrary decision as to what particular items to use as the TBR and irrelevant sets of items. Thus, it is possible that attentional capture by the deviant stimulus in Experiments 1 and 3 had an impact only because that stimulus was a digit and therefore was confused with the digits forming the rehearsal cohort. Of course, if a different category of items had been presented in the two sequences, the question would then have been whether or not the fact that the identity of the irrelevant deviant was categorically incongruent with the TBR set mediated the disruption. We suspect that the relationship between the two sequences in terms of categorical identity was not an influential factor. Nevertheless, to check on this, in Experiment 4 we reverted to the design of Experiment 1 and sought to replicate the deviation effect in a task in which the TBR list consisted of consonants and the irrelevant items were digits. Moreover, we also sought to replicate the observation that the deviation effect is comparable in magnitude regardless of whether it is embedded within an SS or CS irrelevant sequence.

Method

Participants. Thirty-four students at Cardiff University participated in the study for either course credit or a small honorarium. All were native English speakers and had normal or corrected-to-normal vision and normal hearing. None had participated in Experiments 1–3.

Apparatus and materials, design, and procedure. All these aspects of the methodology were identical to Experiment 1, except that the TBR sequences were constructed from random orderings of the consonants *B, H, J, K, L, M, Q,* and *S*. Also, this list (in this order) was written at the top of each page of the response booklet. This was done to bring the primary task more in line with the demand of that in Experiment 1 where a well-known closed set of items was used (*1–9*), which makes the task one of order not item memory.

Results and Discussion

Figure 5 shows the mean percentage of items correctly recalled in the five auditory conditions across the eight serial positions in Experiment 4. The overall pattern across the four irrelevant sound conditions is the same as that found in Experiment 1: There is again a classical CS effect, but most importantly, performance is poorer in the SS + ISI *d* condition ($M = 54.51$, $SE = 2.90$) than in the SS condition ($M = 56.72$, $SE = 3.13$) and poorer in the CS + ISI *d* condition ($M = 49.83$, $SE = 2.86$) than in the CS condition ($M = 52.57$, $SE = 2.63$).

An initial repeated-measures ANOVA revealed a main effect of auditory condition, $F(4, 132) = 13.31$, $MSE = 6.43$, $p < .001$, and planned contrasts showed that all irrelevant sound conditions except the SS condition produced reliably poorer performance than quiet (all ps $< .05$). When we repeated the ANOVA with only the four key conditions of interest (the four irrelevant sound conditions), we found a main effect of state, $F(1, 33) = 16.42$, $MSE = 7.25$, $p < .001$, a main effect of deviation, $F(1, 33) = 7.40$, $MSE = 5.07$, $p < .01$, and a main effect of serial position, $F(7, 231) = 38.35$, $MSE = 11.70$, $p < .001$. The only significant interaction was Deviation × Serial Position, $F(7, 231) = 2.63$, $MSE = 2.51$, $p < .05$. On inspection of Figure 5, this interaction may have emerged because the deviation effect was somewhat more compelling at the early part of the serial position curve. However, given that no such interaction was observed in Experiments 1 and 3, and given the novelty of the deviation effect, we feel it would be premature to attribute any theoretical significance to this interaction. It is important to note that there was again no indication of a State × Deviation interaction, $F(1, 33) = 0.086$, $MSE = 5.17$, $p = .77$ (Cohen's $d = -0.102$; again, power to detect a medium effect size was relatively low [.29], however, the nonsignificant trend is for the deviation effect to be larger in the CS context).

In sum, the deviation effect obtained in Experiment 1 was replicated when the TBR items (consonants) were taken from a category different from that of the irrelevant tokens (digits). Moreover, the deviation effect was once again apparent, and to a comparable degree, irrespective of whether the deviation occurred within a CS or SS sequence. It is noticeable that the classical CS effect was again not as compelling in this experiment as it was in Experiment 1. One possibility is that the poorer overall level of

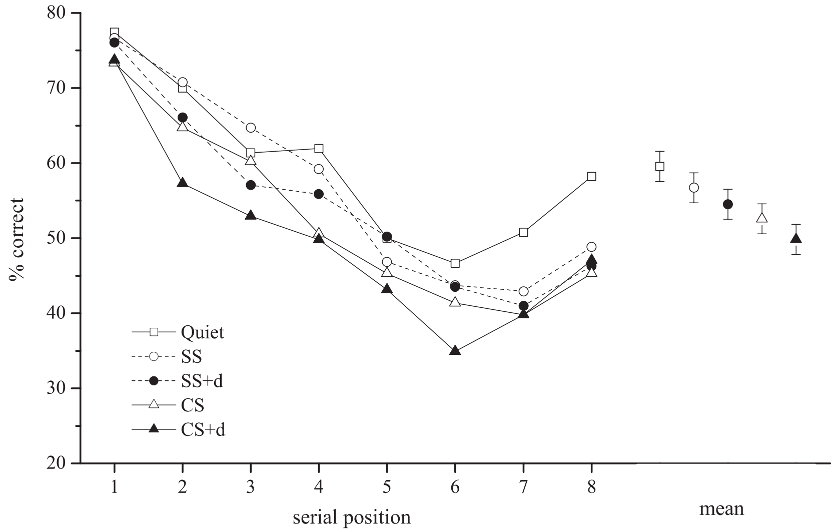

Figure 5. Mean percentage of items correctly recalled in the quiet, steady-state (SS), steady-state + interstimulus interval (ISI) deviation (SS + ISI *d*), changing-state (CS), and changing-state + ISI deviation (CS + ISI *d*) conditions in Experiment 4. Error bars represent 95% confidence intervals.

performance in this experiment compared with Experiment 1 may have rendered the experiment somewhat less sensitive to the CS effect. This general depression in performance may have been due to the change in the TBR set; despite providing the participants access to the set of consonants used, consonants may have been more difficult to recall in order because of relatively less familiarity with this set than with the digit-set *1–9*.

Another possibility is that in Experiments 1 and 2, but not in 4, an order incongruence effect (Hughes & Jones, 2005) might have contributed to the disruption obtained in the CS compared with the SS condition. Hughes and Jones (2005) found that having the irrelevant and relevant items taken from the same set of items increases the degree of disruption compared with when the two sequences are unrelated (TBR digits, irrelevant letters) but only if, in the "identical" condition, the irrelevant and relevant items are presented in incongruent serial orders. Whether or not the deviation effect interacts in some more subtle way with the categorical relationship between the irrelevant and relevant items would certainly be worth pursuing but is beyond the scope of the present article.

It is of course possible that the deviation effect is somehow reliant on the irrelevant sound being made up of speech tokens. Thus, whereas it is well established that the classical CS effect is found with nonspeech stimuli (e.g., D. M. Jones & Macken, 1993), further research is needed to examine whether the same generalization holds for the novel deviation effect demonstrated in the current series. In light of physiological research on the OR, however, in which nonspeech stimuli are typically used, we would not expect speech to play a special role in this effect. Regardless of the foregoing considerations, the key implication of Experiment 4 remains: The deviation effect obtained in Experiment 1 was not reliant on the fact that the irrelevant and relevant items came from the same verbal category.

Practice Analysis

An obvious issue that we have not addressed as yet is whether or not practice over the course of an experiment with sequences containing a temporal deviant diminishes the impact of the deviant. That is, it is conceivable that participants form a higher level neural model that represents the type of sequence to expect across the experimental block (as opposed to a neural model representing the nature and timing of the stimuli presented within a given trial) and that this neural model serves to diminish the novelty—and hence attention-capturing power—of a temporal deviant. In other words, habituation to the deviant may occur at a macroscopic, across-trials, level. It is well established that the classical CS irrelevant sound effect is not subject to habituation across an experimental session (Hellbrück, Kuwano, & Namba, 1996; D. M. Jones, Macken & Mosdell, 1997; Tremblay & Jones, 1998). However, a position we promote later is that the deviation effect is qualitatively different from the classical irrelevant sound effect and on this basis it seems possible that the deviation effect is indeed modulated by the experiment-wide context. We therefore conducted a meta-analysis of the data generated from the present series to examine whether the deviation effect diminishes with practice across an experimental session.[3] We also sought to corroborate previous observations that the classical CS effect does not diminish with practice (D. M. Jones et al., 1997).

Method

Deviation effect. The data from the experimental block for each experiment that produced a deviation effect (Experiments 1, 3, and 4) were divided into fifths (or five subblocks) extending from the 1st fifth to the 5th fifth of trials in the block. By pooling the data across the three experiments for each of the five subblocks, we had data from 16 trials with a deviant (6 SS + ISI *d* trials and 6 CS + ISI *d* trials from Experiments 1 and 4 and 4 SS + ISI *d* trials from Experiment 3) and data from 16 trials without a deviant (6 SS trials and 6 CS trials from Experiments 1 and 4 and 4 SS trials from Experiment 3).

CS effect. For this analysis, the data from the experimental block for each experiment that produced a classical CS effect (Experiments 1, 2, and

[3] Thanks to Alice Healy for pointing out during the review process that it would be worthwhile to conduct this practice analysis.

4) were again divided into five subblocks. By pooling the data across the three experiments for each of the five subblocks, we had data from 9 SS trials and 9 CS trials.

Results and Discussion

Deviation effect. Figure 6A shows the mean percentage of items correctly recalled (collapsed across serial position) on deviant trials compared with no-deviant trials for each of the five subblocks. Although there is some indication that the effect is smaller for the fifth subblock, two sets of statistical analyses failed to reveal any compelling evidence of a diminution in the deviation effect across the five subblocks. First, a 2 (deviant compared with no-deviant trials) × 5 (subblock) repeated-measures ANOVA revealed a main effect of deviation, $F(1, 87) = 20.7$, $MSE = 87.50$, $p < .001$, a main effect of subblock, $F(4, 348) = 6.69$, $p < .001$—reflecting a general practice effect—but, most important, there was no significant Deviation × Subblock interaction, $F(4, 348) = 1.52$, $MSE = 78.34$, $p = .20$.

However, it could be argued that testing whether there was a Deviation × Subblock interaction is not the most appropriate or effective way of examining whether there was a diminution in the deviation effects across the five subblocks. As a second more powerful test, therefore, we conducted a profile analysis to compare the profiles for deviant and no-deviant trials across the five subblocks. With Wilks' criterion, the levels test indicated that recall was significantly higher for no-deviant trials, $F(1, 87) = 20.7$, $p < .001$, and the significant flatness test indicated a general practice effect, $F(4, 84) = 4.41$, $p < .01$. Most important, the profiles did not deviate significantly in terms of parallelism, $F(4, 84) = 1.45$, $p = .225$, which suggests that there was no specific practice effect associated with exposure to deviant trials.

CS effect. Figure 6B shows the mean percentage of items correctly recalled (collapsed across serial position) on SS trials compared with CS trials for each of the five subblocks. Although a general practice effect was again in evidence, there is little indication that the classical CS effect diminishes across an experimental session thereby replicating the results of previous studies (e.g., D. M. Jones et al., 1997). Both the statistical approaches to examining a specific practice effect corroborated this conclusion. First, a 2 (state: SS compared with CS) × 5 (subblock) repeated-measures ANOVA showed a main effect of state, $F(1, 88) = 48.55$, $MSE = 185.23$, $p < .001$, a main effect of subblock, $F(4, 352) = 6.43$, $MSE = 251.90$, $p < .001$, but no State × Subblock interaction, $F(4, 352) = 1.61$, $MSE = 199.05$, $p = .17$. The results of the profile analysis performed to compare SS with CS across the five subblocks corroborated this conclusion. With Wilks' criterion, the levels test showed a CS effect, that is, significantly poorer performance in CS trials than in SS trials, $F(1, 88) = 48.55$, $p < .001$. The flatness test revealed a general practice effect, $F(4, 85) = 4.99$, $p < .001$, but again, there was no specific practice effect associated with exposure to the CS trials as indicated by the results of the parallelism test, $F(4, 85) = 1.42$, $p = .233$.

In summary, we found no statistical evidence for a diminution in the impact of a temporal deviant with practice with deviant trials across an experimental session. It is of course possible that specific practice effects could be observed with a different experimental design. For example, it seems possible that the presence of no-deviant trials interpolated between the deviant trials may have precluded a process of habituation operating at a macroscopic, across-trial level. This could be tested easily by taking a block of trials that contain the (same) deviant event and comparing performance on these trials with a block of trials that do not contain the deviant event. In this case, the effect of the deviant may indeed be shown to diminish across the block of with-deviant trials. The present practice analysis also replicated the observation that the classical CS effect does not diminish across an experimental session (D. M. Jones et al., 1997). Unlike the case with the deviation effect, it has already been established that this null practice effect is found regardless of whether the SS and CS conditions are blocked or are interleaved within a block (D. M. Jones et al., 1997).

General Discussion

The impact of the series may be summarized as follows: We showed a deviation effect within the context of a short-term memory task for the first time; this effect occurs to a roughly similar extent in sequences with changing items and in sequences with a single repeating item (Experiments 1 and 4). Moreover, some types of cognitive activity—arguably ones that represent postencoding processes—seem immune to a deviation effect (Experiment 2). The impact of the deviant does not seem to be reliant on the presence of a highly predictable temporal relationship between the irrelevant and relevant sequence (Experiment 3). Neither is the effect contingent on the particular relationship between the irrelevant and relevant material in terms of verbal

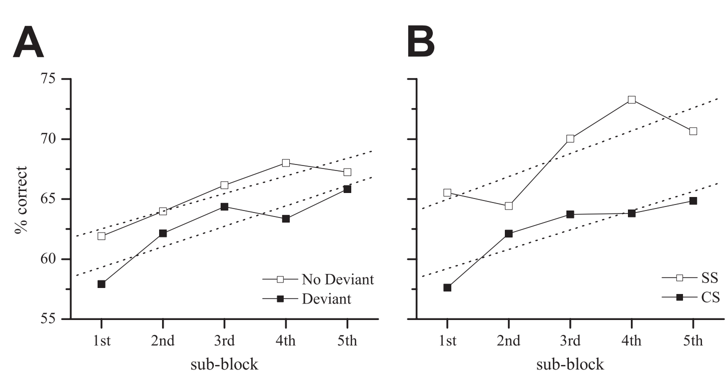

Figure 6. Mean percentage of items correctly recalled for no-deviant and deviant trials (A) and steady-state (SS) and changing-state (CS) trials (B) across five subblocks of three pooled, experimental sessions.

category (Experiment 4). Finally, a practice analysis failed to reveal any compelling evidence that the deviation effect diminishes with exposure to trials containing a deviant. However, it remains possible that this only holds in the context of the present type of design in which there were no-deviant trials interleaved with the deviant trials. Notwithstanding this caveat, the preliminary evidence presented here suggests that the deviation effect is enduring and thus may turn out to have applied as well as theoretical implications (for discussions of the applied implications of the classical CS effect, see Hughes & Jones, 2001, 2003b).

The first, general, implication of the present results is that they indicate that a deviant auditory stimulus seems to capture attention even when it is embedded in a stream of information that is completely irrelevant to the primary task. Although this has been demonstrated before in the context of physiological research (see, e.g., Näätänen, 1992), in behavioral studies the deviation typically occurs at some (unpredictable) point within the same sequence that contains the task-relevant information (Dalton & Lavie, 2004) if not within the same stimulus that acts as the target (e.g., Mondor et al., 1998; Schröger & Wolff, 1998). Second, the precise pattern of results helps to adjudicate between two approaches to the phenomenon of attentional capture. The pattern does not favor the idea that attentional capture can be construed as a result of a stimulus violating an aggregate model of the preceding stimuli. If this were the case, the CS stimuli would have already had the power to capture attention (cf. Cowan, 1995), and it seems reasonable to assume that any additional deviation (from the regular ISI) would not have had as much impact as when it occurred within the context of a highly habituated SS sequence. In contrast, the pattern of results is consistent with the notion that attentional capture is the product of a departure from an algorithm-based model. On this account, no item within either a standard SS or a standard CS sequence would violate any rule and therefore none of the items assume their disruptive potency by capturing attention. Thus, when a stimulus is introduced that does deviate from a rule to which the sequence conforms (e.g., "the ISI is regular"), that deviation would be functionally identical in terms of attention-capturing power irrespective of whether it is embedded in an SS or CS sequence.

In providing support for an algorithm- over an aggregate-model account of attentional capture, the pattern of results also undermines the suggestion by Cowan (1995) that an SS sequence has less impact on serial recall because none of the items depart from some aggregate neural model. The finding that attentional capture seems only to disrupt recall when the deviation occurs during the presentation of the TBR items further bolsters the view that it is functionally distinct from the classical CS effect. Indeed, in light of the present findings and in convergence with several other observations in the literature, we would argue that an attentional capture-based approach to the irrelevant sound effect is unlikely to prove fruitful.

First, if attentional capture is a key determinant of the irrelevant sound effect, then the impact of the irrelevant sound should diminish over many trials during an experiment. This is because such repeated exposure to the stimuli should allow for the fabrication of an aggregate-based model of the stimuli so that attentional capture should become less likely over time. However, as we have already discussed, the degree of disruption does not diminish over the course of an experiment even when the SS and CS conditions are blocked (Hellbrück et al., 1996; D. M. Jones et al., 1997; Tremblay & Jones, 1998; but see Banbury & Berry, 1997) nor between subsequent experimental sessions in which the same stimuli are used to make up the irrelevant sequences (Ellermeier & Zimmer, 1997; Hellbrück et al., 1996).

Second, the account would not predict the *token-dose effect* (Bridges & Jones, 1996). This refers to the finding that as the number of different tokens per unit time within a trial increases, so too does the degree of disruption. If attentional capture becomes less likely the more times a stimulus is presented (resulting from the increasingly better specified aggregate-based neural model), then performance should in fact be better the higher the token dose (or certainly not worse).

Third, following a similar logic to that with the token-dose effect, the higher the number of different tokens used in the irrelevant stream within a trial (*token set-size*) the more likely it is that attentional capture will occur. That is, with a low token-set size (e.g., *a, b, a, b, a, b* . . .) each token would be repeated relatively more often than with a higher token-set size (e.g., *a, b, c, d, a, b* . . .), thus serving to fashion a more accurate aggregation model more quickly and, in turn, leading to a diminishing likelihood of capture. Thus, the attentional capture account of the CS effect would predict a monotonic increase in disruption as the token-set size increases. However, a change between immediately successive tokens (low token-set size) is sufficient to produce a disruptive effect, with the addition of further tokens (higher token-set size), having no further reliable effect on disruption (Hughes & Jones, 2005; Tremblay & Jones, 1998; but see Campbell, Beaman, & Berry, 2002).

Fourth, the attentional capture account is too general insofar as it cannot account for the fact that tasks containing a seriation component are particularly vulnerable to the CS irrelevant sound effect (Beaman & Jones, 1997, 1998; Jones & Macken, 1993). Finally, it has been shown that, up to a point, the larger the differences between immediately successive items the greater the disruption to recall. However, when successive auditory items are segregated into separate streams as a result of particularly large differences between them (e.g., in terms of pitch: D. M. Jones, Alford, Bridges, Tremblay, & Macken, 1999; or in terms of location: D. M. Jones & Macken, 1995) so that each stream contains one unvarying item, disruption is attenuated. There seems no obvious reason, on the attentional capture account, why there should be a nonmonotonic relationship between the degree of CS and its disruptive power.

This latter observation demonstrates that a critical characteristic of the CS effect is that only variation superimposed on a common substrate (i.e., changes within the parameters of a single coherent stream) produces the effect. That is, to be highly disruptive of serial recall, the successive stimuli must exhibit acoustic change but must still at the same time share a more fundamental underlying acoustic similarity (e.g., a sequence of different words spoken in the same voice; see D. M. Jones & Macken, 1995; D. M. Jones, Alford, et al., 1999). This observation may hold the key to understanding why items in a CS sequence do not exert their disruptive power by capturing attention from the primary task. In abstract terms, attentional capture—both by auditory and visual stimuli (for an overview, see Eimer et al., 1996)—is typically associated with the detection of the onset of a new object that potentially warrants further scrutiny because of its potential danger or opportunity for the organism (e.g., Johnston & Strayer, 2001). However, variation on a common ground—the precondition for a marked CS effect (e.g., D. M. Jones, Alford, et al., 1999)—does not, by definition, signal the onset of a new object; rather, the changes are perceived as fluctuations within the properties of a single event or temporally extended object over time.

It has been argued that the CS effect is better accounted for in terms of a conflict between two similar seriation processes, one involved in the attentionally driven rehearsal of the TBR items and the other a by-product of preattentively organizing the component elements of the irrelevant sound into a single coherent stream. More specifically, the process of integrating elements into the same stream despite their exhibiting some acoustic differences gives rise to information pertaining to the order of the elements, which, in turn, corrupts the rehearsal process (Hughes & Jones, 2003a; D. M. Jones, 1993; D. M. Jones, Beaman, & Macken, 1996; Tremblay & Jones, 1998). When the difference between successive elements is particularly large, the elements are partitioned into separate streams, information pertaining to their order is impoverished, and disruption to recall is correspondingly attenuated (e.g., D. M. Jones, Alford, et al., 1999).

On this interference-by-process account, in stark contrast to the attentional capture account (Cowan, 1995), it is precisely because each acoustically different element is not perceived as deviating from an algorithm governing the sequential integration of the elements into a stream that gives rise to order information and therefore disruption of serial recall. In other words, the CS effect will be strongest just below the threshold at which the items will be perceived as different enough from one another to be perceived as novel and therefore potentially attention-capturing events.

The foregoing analysis raises the intriguing possibility that it is the algorithmic process of auditory streaming that ensures that only novel auditory events are endowed with the capacity to capture attention. Indeed, it would be dysfunctional for changes in external input to capture attention if those changes merely constituted variations in energy arising from the same auditory source. Auditory streaming therefore may play a key role in mediating the requirement that attentional selection can be engaged on current goals while simultaneously remaining open to interruption from potentially important information deriving from the previously unattended portion of the auditory scene (Allport, 1989; Hughes & Jones, 2003b). That is, attentional capture may be invoked only by a departure from some prevailing rule deployed in the preattentive organization of irrelevant material into streams, not from a rule that characterizes the stimulus sequence per se.[4] Certainly, the present data are not definitive in this regard, only suggestive. Whether the algorithm on which capture is based is stream-based or external stimulus-based is a hypothesis that clearly merits further investigation.

[4] To illustrate the possible distinction between a violation of a perceptual stream and a violation of the sequence per se, consider the situation in which the deviation is a repetition of an item presented within a male-spoken sequence (e.g., *5 2 4 1 6 8 9 9 3 7*). Let us assume that the second "*9*" produces a deviation effect (i.e., captures attention and thereby disrupts serial recall). If the algorithmic model is stream-based rather than sequence-based, then the second "*9*" should be stripped of its attention-capturing power if alternate items are presented such that they are likely to be partitioned perceptually into two distinct streams. This could be achieved by presenting the items in an alternating male-voice–female-voice fashion (e.g., *5* **2** *4* **1** *6* **8** *9* **9** *3* **7**; female-spoken items shown in bold). That is, because the second "**9**" would no longer be perceived as part of the same stream as that which contains the first (male-spoken) "*9*," the repetition may no longer be perceived and, in turn, no attentional capture effect should be evident. If, in contrast, the algorithmic model is sequence-based, the repetition should capture attention regardless of how the stimuli are perceptually organized. This hypothesis is currently under investigation in our laboratory in Cardiff.

References

Allport, D. A. (1989). Visual attention. In M. I. Posner (Ed.), *Foundations of cognitive science* (pp. 631–682). Cambridge, MA: MIT.

Banbury, S., & Berry, D. C. (1997). Habituation and dishabituation to speech and office noise. *Journal of Experimental Psychology: Applied, 3,* 181–195.

Beaman, C. P., & Jones, D. M. (1997). The role of serial order in the irrelevant speech effect: Tests of the changing state hypothesis. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23,* 459–471.

Beaman, C. P., & Jones, D. M. (1998). Irrelevant sound disrupts order information in free as in serial recall. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 51*(A), 615–636.

Bregman, A. S. (1990). *Auditory scene analysis: The perceptual organization of sound.* Cambridge, MA: MIT.

Bridges, A. M., & Jones, D. M. (1996). Word dose in the disruption of serial recall by irrelevant speech: Phonological confusions or changing-state? *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 49*(A), 919–939.

Campbell, T., Beaman, C. P., & Berry, D. C. (2002). Auditory memory and the irrelevant sound effect: Further evidence for changing-state disruption. *Memory, 10,* 199–214.

Cohen, J. D., MacWhinney, B., Flatt, M., & Provost, J. (1993). PsyScope: A new graphic interactive environment for designing psychology experiments. *Behavioral Research Methods, Instruments & Computers, 25,* 256–271.

Colle, H. A., & Welsh, A. (1976). Acoustic masking in primary memory. *Journal of Verbal Learning and Verbal Behavior, 15,* 17–32.

Cowan, N. (1995). *Attention and memory: An integrated framework.* London, England: Oxford University Press.

Dalton, P., & Lavie, N. (2004). Auditory attentional capture: Effects of singleton distractor sounds. *Journal of Experimental Psychology: Human Perception and Performance, 30,* 180–193.

Eimer, M., Nattkemper, D., Schröger, E., & Prinz, W. (1996). Involuntary attention. In O. Neumann & A. F. Sanders (Eds.), *Handbook of perception and action, Vol. 3* (pp. 389–446). London: Academic Press.

Ellermeier, W., & Zimmer, K. (1997). Individual differences in the susceptibility to the "irrelevant speech effect." *Journal of the Acoustical Society of America, 102,* 2191–2199.

Elliott, E. M. (2002). The irrelevant-speech effect and children: Theoretical implications of developmental change. *Memory & Cognition, 30,* 478–487.

Hellbrück, J., Kuwano, S., & Namba, S. (1996). Irrelevant background speech and human performance: Is there long-term habituation? *Journal of the Acoustical Society of Japan, 17,* 239–247.

Hughes, R. W., & Jones, D. M. (2001). The intrusiveness of sound: Laboratory findings and their implications for noise abatement. *Noise & Health, 4,* 55–74.

Hughes, R. W., & Jones, D. M. (2003a). A negative order-repetition priming effect: Inhibition of order in unattended auditory sequences? *Journal of Experimental Psychology: Human Perception and Performance, 29,* 199–218.

Hughes, R. W., & Jones, D. M. (2003b). Indispensable benefits and unavoidable costs of unattended sound for cognitive functioning. *Noise & Health, 6,* 63–76.

Hughes, R. W., & Jones, D. M. (2005). The impact of order incongruence between a task-irrelevant auditory sequence and a task-relevant visual sequence. *Journal of Experimental Psychology: Human Perception and Performance, 31,* 316–327.

Johnston, W. A., & Strayer, D. L. (2001). A dynamic, evolutionary perspective on attentional capture. In C. Folk & B. Gibson (Eds.), *Attraction, distraction, and action: Multiple perspectives on attentional capture* (pp. 375–397). New York: Elsevier.

Jones, D. M. (1993). Objects, streams and threads of auditory attention. In A. D. Baddeley & L. Weiskrantz (Eds.), *Attention: Selection, awareness, and control,* (pp. 167–198). Oxford, England: Clarendon Press.

Jones, D. M., Alford, D., Bridges, A., Tremblay, S., & Macken, W. J. (1999). Organizational factors in selective attention: The interplay of acoustic distinctiveness and auditory streaming in the irrelevant sound effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25,* 464–473.

Jones, D. M., Beaman, C. P., & Macken, W. J. (1996). The object-oriented episodic record model. In S. Gathercole (Ed.), *Models of short-term memory* (pp. 209–238). Mahwah, NJ: Erlbaum.

Jones, D. M., & Macken, W. J. (1993). Irrelevant tones produce an irrelevant speech effect: Implications for phonological coding in working memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19,* 369–381.

Jones, D. M., & Macken, W. J. (1995). Organizational factors in the effect of irrelevant speech: The role of spatial location and timing. *Memory & Cognition, 21,* 318–328.

Jones, D. M., Macken, W. J., & Mosdell, N. (1997). The role of habituation in the disruption of recall performance by irrelevant sound. *British Journal of Psychology, 88,* 549–564.

Jones, D. M., Madden, C., & Miles, C. (1992). Privileged access by irrelevant speech to short-term memory: The role of changing state. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 44*(A), 645–669.

Jones, M. R. (1976). Time, our lost dimension: Toward a new theory of perception, attention, and memory. *Psychological Review, 83,* 323–335.

Jones, M. R. (2001). Temporal expectancies, capture, and timing in auditory sequences. In C. Folk & B. Gibson (Eds.), *Attraction, distraction, and action: Multiple perspectives on attentional capture* (pp. 191–229). New York: Elsevier.

Lashley, K. S. (1951). The problem of serial order in behavior. In L. A. Jeffress (Ed.), *Cerebral mechanisms in behavior: The Hixon symposium* (pp. 112–146). New York: Wiley.

Macken, W. J., Mosdell, N., & Jones, D. M. (1999). Explaining the irrelevant sound effect: Temporal distinctiveness or changing state? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25,* 810–814.

Miles, C., Jones, D. M., & Madden, C. A. (1991). Locus of the irrelevant speech effect in short-term memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17,* 578–584.

Mondor, T. A., Zatorre, R. J., & Terrio, N. A. (1998). Constraints on the selection of auditory information. *Journal of Experimental Psychology: Human Perception and Performance, 24, 66*–79.

Näätänen, R. (1992). *Attention and brain function.* Mahwah, NJ: Erlbaum.

Nicholls, A. P., & Jones, D. M. (2002). Perceptual organization in short-term memory: The sandwich effect reassessed. *Memory & Cognition, 30,* 81–88.

Nordby, H., Roth, W. T., & Pfefferbaum, A. (1988). Event-related potentials to breaks in sequences of alternating pitches or interstimulus intervals. *Psychophysiology, 25,* 262–268.

Saarinen, J., Paavilainen, P., Schröger, E., Tervaniemi, M., & Näätänen, R. (1992). Representation of abstract attributes of auditory stimuli in the human brain. *NeuroReport, 3,* 1149–1151.

Salamé, P., & Baddeley, A. (1982). Disruption of short-term memory by unattended speech: Implications for the structure of working memory. *Journal of Verbal Learning and Verbal Behavior, 21,* 150–164.

Salamé, P., & Baddeley, A. D. (1989). Effects of background music on phonological short-term memory. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 41*(A), 107–122.

Schröger, E. (1997). On the detection of auditory deviants: A preattentive activation model. *Psychophysiology, 34,* 245–257.

Schröger, E., & Wolff, C. (1996). Mismatch response to changes in sound location. *NeuroReport, 7,* 3005–3008.

Schröger, E., & Wolff, C. (1998). Behavioral and electrophysiological effects of task-irrelevant sound change: A new distraction paradigm. *Cognitive Brain Research, 7,* 71–87.

Seger, C. A. (1994). Implicit learning. *Psychological Bulletin, 115,* 163–196.

Sokolov, E. N. (1963). *Perception and the conditioned reflex.* London: Pergamon Press.

Sokolov, E. N. (1975). The neuronal mechanisms of the orienting reflex. In E. N. Sokolov & O. S. Vinogradova (Eds.), *Neuronal mechanisms of the orienting reflex* (pp. 217–235). Hillsdale, NJ: Erlbaum.

Stroop, J. R. (1935). Interference in serial verbal reactions. *Journal of Experimental Psychology, 18,* 643–661.

Tipper, S. P., Howard, L. A., & Jackson, S. R. (1997). Selective reaching to grasp: Evidence for distractor interference effects. *Visual Cognition, 4,* 1–38.

Tremblay, S., & Jones, D. M. (1998). Role of habituation in the irrelevant sound effect: Evidence from the effects of token set size and rate of transition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24,* 659–671.

Unger, S. M. (1964). Habituation of the vasoconstrictive orienting reaction. *Journal of Experimental Psychology, 67,* 11–18.

Velden, M. (1978). Some necessary revisions of the neuronal model concept of the orienting response. *Psychophysiology, 15,* 181–185.

Woods, D. L. (1990). The physiological basis of selective attention: Implications of event-related potential studies. In J. W. Rohrbaugh, R. Parasuraman, & R. Johnson, Jr. (Eds.), *Event-related brain potentials: Basic issues and applications* (pp. 178–209). New York: Oxford University Press.

Yantis, S. (1998). Control of visual attention. In H. Pashler (Ed.), *Attention.* Hove, England: Psychology Press.

Received May 10, 2004
Revision received January 5, 2005
Accepted January 6, 2005 ■