

Directed Attention Eliminates ‘Change Deafness’ in Complex Auditory Scenes

Ranmalee Eramudugolla,¹ Dexter R.F. Irvine,¹
Ken I. McAnally,² Russell L. Martin,²

and Jason B. Mattingley^{3,*}

¹Department of Psychology

Monash University

Victoria 3800

Australia

²Defence Science and Technology Organisation

Victoria 3207

Australia

³Cognitive Neuroscience Laboratory

School of Behavioural Science

University of Melbourne

Victoria 3010

Australia

Summary

In natural environments that contain multiple sound sources, acoustic energy arising from the different sources sums to produce a single complex waveform at each of the listener’s ears. The auditory system must segregate this waveform into distinct streams to permit identification of the objects from which the signals emanate [1]. Although the processes involved in stream segregation are now reasonably well understood [1–3], little is known about the nature of our perception of complex auditory scenes. Here, we examined complex scene perception by having listeners detect a discrete change to an auditory scene comprising multiple concurrent naturalistic sounds. We found that listeners were remarkably poor at detecting the disappearance of an individual auditory object when listening to scenes containing more than four objects, but they performed near perfectly when their attention was directed to the identity of a potential change. In the absence of directed attention, this “change deafness” [4] was greater for objects arising from a common location in space than for objects separated in azimuth. Change deafness was also observed for changes in object location, suggesting that it may reflect a general effect of the dependence of human auditory perception on attention.

Results and Discussion

Mechanisms of attention are crucial for selecting information from the different sense modalities for further processing. In complex environments in which multiple simultaneous stimuli compete for selection, attentional mechanisms enhance the processing of relevant sensory inputs and attenuate or suppress those that are irrelevant [5, 6]. In some cases, filtering of irrelevant inputs can be so strong that unattended stimuli fail to reach awareness altogether [7]. Within the auditory domain, previous studies have shown that when attention

is focused on a particular sound stream, listeners do not have a full and detailed awareness of other features of their auditory environment. For example, unexpected but suprathreshold changes in a concurrent stream often go unnoticed [7], as do unattended features of a monitored stream [4]. These observations seem contrary to our subjective experience of being fully aware of our immediate auditory environment: In the absence of competing demands on attention, we would expect that salient changes in the auditory environment, for example the disappearance of a person’s voice during a conversation or a sudden shift in the location of a siren, would be relatively easy to detect.

Here, we used a novel change-detection task to examine the role of selective attention in the perception of complex scenes comprising multiple naturalistic sounds that were presented concurrently. Artificial auditory scenes comprising multiple sounds (or “auditory objects”) were created within virtual auditory space (VAS) [8]. VAS stimuli are perceived as originating outside the head and at distinct spatial locations (see Figure 1), and they can be localized with comparable accuracy to external stimuli [9, 10]. Participants listened to auditory scenes containing four, six, or eight objects and had to detect a salient change (viz., the disappearance of one of the objects or a switch in the locations of two of the objects) with or without the benefit of an attentional cue.

In a typical trial of the object-disappearance task, listeners heard two versions of an auditory scene, each 5 s long, from the second of which one object was missing in test trials. In the “directed-attention” condition (Figure 1B), participants were shown the name of an object (e.g., “cello”) on the computer screen and had to determine whether that object was missing from the second version. In the “nondirected-attention” condition (Figure 1A), participants had to attend to all objects in the initial version of the scene and determine whether one of the objects was missing in the second version. The two versions of the scene were separated by a 500 ms burst of white noise (Figure 1) to delimit the two versions and to mask any transients or echoic memory trace that might cue the listener’s attention to the change.

The directed- and nondirected-attention conditions were presented in separate blocks; in each block, a change occurred in 75% of trials, and the remaining 25% were catch trials, in which the two scenes were identical. In each trial, participants made a yes/no judgment about whether a change had occurred. Two measures of change detection were calculated: the percentage of changes detected and perceptual sensitivity (d'). A measure of response criterion (C), which is independent of d' [11], was also calculated.

In experiment 1, we used a standard set of Head Related Transfer Functions (HRTFs—see [Experimental Procedures](#)) to create the auditory scenes. Twenty-eight participants were tested. In the directed-attention condition, participants’ ability to report the disappearance of the cued object was nearly perfect for all scene

*Correspondence: j.mattingley@psych.unimelb.edu.au

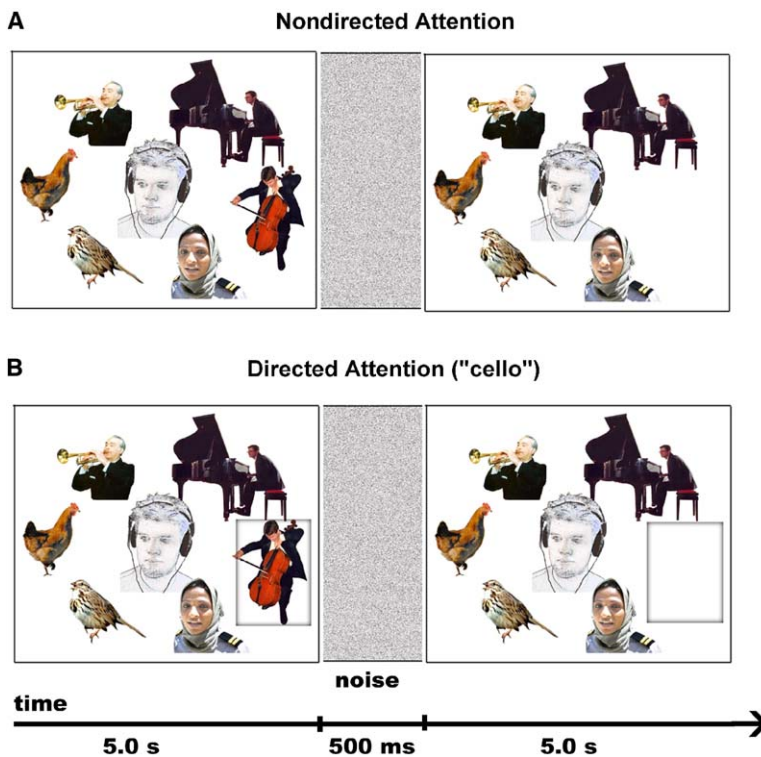


Figure 1. Schematic of a VAS Scene Containing Six Auditory Objects

The icons around the listener's head represent the objects (viz., trumpet reveille, piano solo, cello solo, female voice, bird's chirrups, and hen's clucking) and the azimuthal locations at which they were generated. Each trial consisted of a 5 s segment of the scene, then a white-noise burst for 500 ms, and then a further 5 s segment of the scene. (A) In the nondirected condition, attention was not cued to any object in the scene. Participants indicated whether any object (the cello in this example) had disappeared from the scene.

(B) In the directed condition, attention was cued to one object (indicated by the box), and participants determined whether that object disappeared (experiments 1 and 2) or exchanged location with another object (experiment 3).

sizes (see Figures 2A and 2B). Thus, despite the complexity of the scenes, participants were able to segregate, identify, and monitor an individual object, provided they could focus their attention on that object. In contrast, change detection in the nondirected-attention condition was remarkably poor: The proportion of changes missed increased with scene size and approached 50% for scenes comprising eight objects (Figure 2A). These differences between the conditions were also evident for the d' measure of sensitivity (Figure 2B). An ANOVA on d' values revealed significantly poorer change detection in the nondirected- than in the directed-attention condition ($F_{1,27} = 129.71$, $p < 0.001$) and a significant decrement in performance with increasing auditory-scene size ($F_{2,41} = 65.00$, $p < 0.001$). Crucially, there was also an interaction between attention condition and scene size ($F_{2,54} = 36.76$, $p < 0.001$): Change detection became significantly poorer as scene size increased in the nondirected-attention condition ($F_{2,54} = 75.46$, $p < 0.001$), but there was no such effect in the directed-attention condition ($F_{2,54} = 2.71$, $p > 0.05$).

These results indicate that when attention is not directed toward an auditory object within a complex scene, explicit detection of a change is remarkably difficult, even when the listener is aware that a change is likely to occur. When attention is directed to the identity of the changed object, detection is independent of the number of objects in a scene over the range tested; but when attention is not so directed, detection deteriorates with increasing scene size. Almost identical results were obtained when eight different participants were tested with individualized HRTFs (see the Supplemental Data available with this article online), indicating

that the change-deafness effect observed in experiment 1 cannot be attributed to the fidelity of the HRTFs.

The sounds associated with the auditory objects used in experiment 1 had substantially overlapping frequency spectra (see Supplemental Data), indicating that performance in the directed-attention condition could not have been based simply on attention to a specific frequency ("listening" [12] or "attention" [13]) band. An alternative possibility is that participants attended to the spatial location of the cued object [14, 15] and responded to a change in that location. To examine this issue, we presented the same auditory scenes in two spatial-separation conditions in experiment 2. In the different-locations condition, each object in the scene was assigned a distinct spatial position in the azimuthal plane, as in experiment 1. In the same-location condition, all objects were assigned to the same spatial location in the azimuthal plane. The "same" locations varied across trials over the range used for different objects in the different-locations condition. Attention was manipulated as in experiment 1, and the same proportion of change to no-change trials was used. Scene size (four, six, or eight sounds) was varied randomly within each block of trials, and all auditory scenes were generated with individualized HRTFs.

As shown in Figures 3A and 3B, elimination of spatial separation between objects in the auditory scenes did not affect change detection in the directed-attention condition. For the nondirected-attention condition (Figures 3D and 3E), however, it resulted in a small but reliable decrease in change-detection performance. The poorer change detection for same-location scenes in the nondirected-attention condition was confirmed sta-

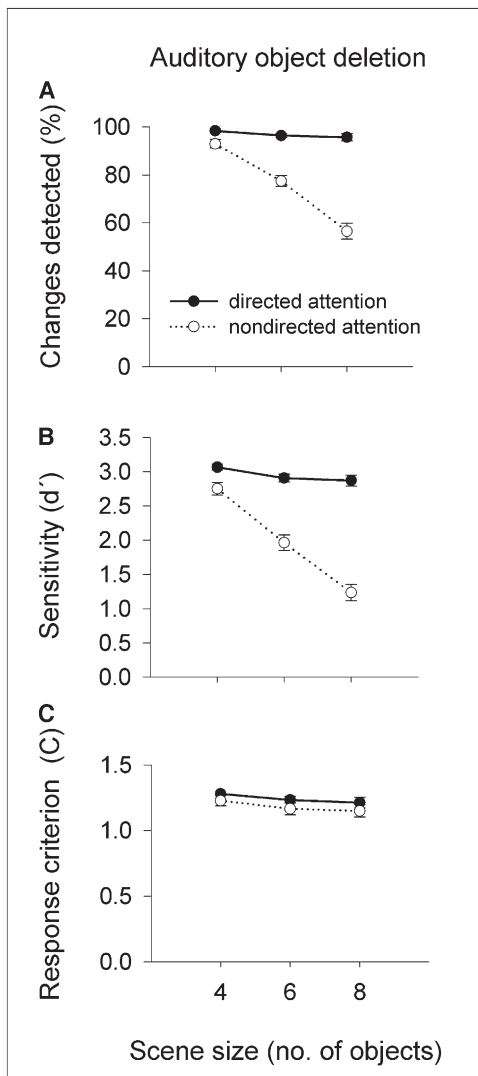


Figure 2. Effect of Attention on Mean Change-Detection Measures (± 1 standard error), as a Function of Auditory-Scene Size

(A) Percentage of changes detected.

(B) Sensitivity (d').

(C) Response criterion. Participants' response criterion did not differ for the two attention conditions and also did not vary with scene size. An ANOVA on the criterion values confirmed this pattern: There was no effect of attention condition ($F_{1,27} = 2.11$, $p > 0.10$) or scene size ($F_{1,45} = 3.22$, $p > 0.05$) and no significant interaction between these factors ($F_{1,39} < 1.0$).

tistically for the d' values ($F_{1,11} = 9.32$, $p < 0.05$). Consistent with the findings from experiment 1, performance also deteriorated significantly with increasing scene size ($F_{2,22} = 28.75$, $p < 0.001$). There was no interaction between scene size and spatial separation ($F_{2,22} = 1.37$, $p > 0.10$).

The fact that listeners performed at near-ceiling levels in the directed-attention condition, even when objects were not spatially separated, suggests that spatial cues played a relatively minor role in auditory streaming and attention to objects in our task. This might, in part, reflect the fact that the auditory objects

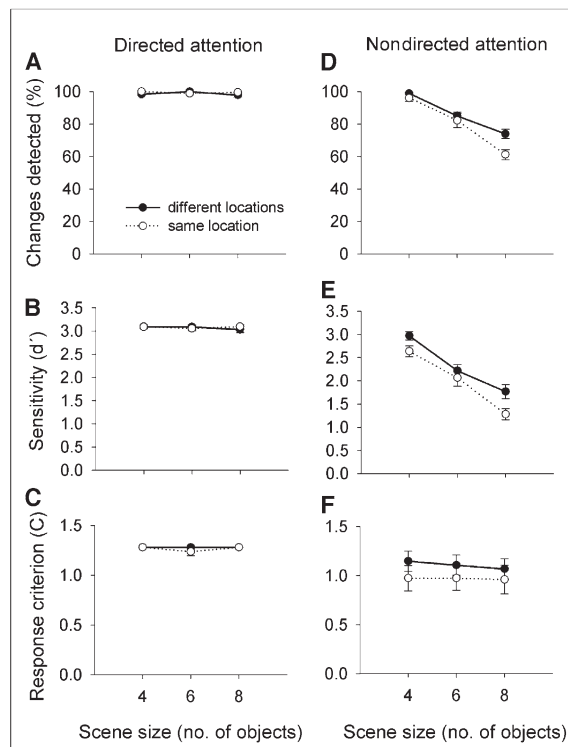


Figure 3. Effect of Spatial Separation of Auditory Objects on Change-Detection Performance

(A–C) Directed-attention condition; (D–F) nondirected-attention condition.

(A and D) Percentage of changes detected.

(B and E) Sensitivity (d'). In the directed-attention condition, d' did not vary with spatial separation ($F_{1,11} < 1.0$) or scene size ($F_{2,22} < 1.0$). In contrast, for the nondirected-attention condition, d' was significantly poorer for the same-location than the different-location condition.

(C and F) Response criterion. Analysis of response criterion in the directed-attention condition revealed no significant effect of spatial separation ($F_{1,11} < 1.0$) or scene size ($F_{2,22} = 2.66$, $p > 0.05$). In the nondirected-attention condition, criterion was significantly different for the two spatial-separation conditions ($F_{2,18} = 16.64$, $p < 0.01$) but did not vary with scene size ($F_{2,22} < 1.0$).

used in the scenes were selected to be perceptually distinctive, with the consequence that listeners paid greater attention to the marked spectrotemporal differences between objects (see [Supplemental Data](#)) than to spatial location.

Previous studies have suggested that listeners may have particular difficulty noticing the disappearance of a single sound stream from a mixture [16] and that the offset of a visual stimulus is less effective in capturing attention than the onset of a stimulus [17]. Thus, the results obtained in experiments 1 and 2 might reflect mechanisms unique to the case of object disappearance. In experiment 3, we therefore investigated the effect of selective attention on listeners' perception of changes in the spatial location of objects in complex scenes. If the change-deafness effect reflects a general limitation in listeners' capacity to fully perceive a complex auditory scene, then it should also be apparent for changes in object location.

The basic paradigm in experiment 3 was identical to that employed in experiment 1 (see Figure 1), except that two objects exchanged locations from the first to the second version of the scene. The minimum spatial separation between objects that exchanged locations was 40°, and although each location change involved two objects in the scene, participants were only required to attend to and report changes in the position of one object. Twenty-six participants, who were also involved in experiment 1, participated in experiment 3. The order of presentation of the two experiments was alternated across participants. As in experiment 1, the auditory scenes were generated with a standard set of HRTFs. In the nondirected-attention condition, participants were required to report whether any object changed location. In the directed-attention condition, participants were given the name of one object and had to report whether that object changed location.

The pattern of results for object-location change was similar to that obtained for object disappearance in experiment 1. As indicated in Figure 4A, participants' detection of location changes was higher in the directed- than in the nondirected-attention condition. The change-detection rate also decreased with increasing scene size in the nondirected-attention condition but remained stable when attention was directed to the change object. Analysis of d' data (Figure 4B) revealed significant main effects of attention condition ($F_{1,25} = 33.95, p < 0.001$) and scene size ($F_{2,50} = 9.64, p < 0.001$) and a significant interaction ($F_{2,50} = 7.85, p < 0.01$). Sensitivity decreased with increasing scene size in the nondirected condition ($F_{2,50} = 12.43, p < 0.001$) but was unaltered across scene size in the directed-attention condition ($F_{2,50} = 1.51, p > 0.50$). These results indicate that unless listeners selectively attend to an object in a complex auditory scene, they are less likely to detect a change in its location as scene complexity increases. The fact that listeners had to attend to spatial aspects of the auditory scene to detect changes in object location suggests that the fidelity of the HRTFs used in this experiment would have influenced the overall level of performance. Results from a separate group of seven participants tested with individualized HRTFs indicated that although individualized HRTFs conferred an advantage in detecting location changes, the overall pattern of results was unchanged (see Supplemental Data).

In the directed-attention condition, detection of location changes (Figures 4A and 4B) was poorer overall than detection of object disappearance (experiment 1; Figures 1A and 1B), suggesting that the former task is more difficult and that source location was not the primary feature used to detect object disappearance in experiment 1. Crucially, however, our findings indicate that detection of *both* types of auditory change is significantly compromised in the absence of directed attention. The response-criterion data for experiment 3 also differed from those obtained in experiment 1. As shown in Figure 4, change detection became more difficult in the absence of directed attention and with increases in scene size, and participants tended to adopt a more liberal response criterion.

The findings from our auditory change-detection tasks reveal a striking limitation in the number of audi-

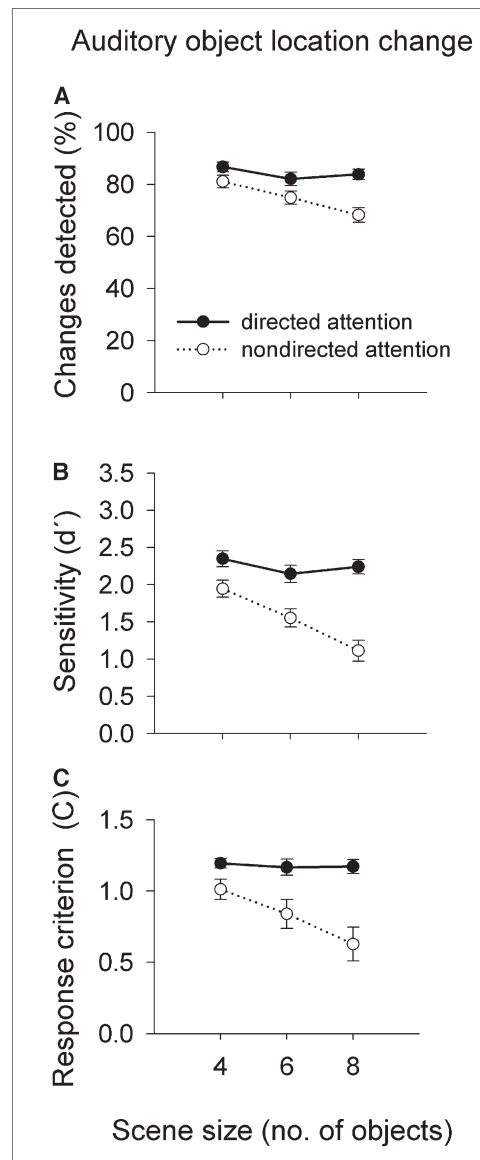


Figure 4. Effect of Attention on the Detection of Changes to Object Location

(A) Percentage of changes detected.

(B) Sensitivity (d').

(C) Response criterion. Analysis of response criterion for changes in object location revealed a main effect of attention condition ($F_{1,25} = 31.63, p < 0.001$) and of scene size ($F_{2,50} = 3.64, p < 0.05$) and an interaction that was marginally significant ($F_{2,40} = 3.17, p = 0.06$).

tory objects that can be monitored concurrently by human listeners. Laboratory studies have shown that humans are extremely skilled at identifying [18] and localizing [9] auditory events that occur in isolation. In the natural world, however, sounds rarely occur alone. Our findings indicate that when a listener's attention is directed to a particular object within a complex auditory scene, the disappearance of that object or a change in its location is rarely missed for scenes containing up to eight different objects. In the absence of

an attentional cue, however, detection of such changes is relatively poor and falls dramatically for auditory scenes containing more than about four objects.

Several studies have examined the role of selective attention in the detection and identification of various types of auditory stimuli [7, 14, 15] and in auditory-scene analysis [3]. However, no previous study has investigated the effect of directed attention on perception of complex auditory scenes comprising naturalistic objects. Our results parallel those from visual change-detection studies, which have shown that normal observers are remarkably insensitive to salient alterations in naturalistic visual scenes across brief interruptions [19–22].

Our observation of change deafness suggests that the human auditory system also relies on attention to detect changes in complex auditory scenes. It is not clear from our data whether the effects of attention on change detection precede or follow the segregation of the separate streams comprising complex scenes. Directed attention might be required for stream segregation per se, in which case the change-deafness effects observed here would reflect a limit in maintaining separate auditory streams in the absence of directed attention. Alternatively, all of the objects in the initial scene might be fully segregated prior to attentional selection. In this case, change deafness would reflect a limit in encoding and storing multiple auditory objects for comparison with a subsequent scene. Whatever the mechanisms, our results indicate that auditory perception is limited by attention and that our experience of a rich and detailed auditory world may be largely illusory.

Experimental Procedures

Auditory Scenes

Each auditory scene was composed of a combination of four, six, or eight sounds drawn at random from the following library of 11 natural sounds presented at equivalent root mean square (RMS) sound pressure levels (70–80 dB): birds chirping, synthesized drum beat, hens clucking, Gregorian chant, piano solo, cello solo, trumpet reveille, male horse-race caller (English), female newsreader (Hindi), police siren, and alarm-clock ring. The sounds had different spectrotemporal patterns, but the frequency spectra of the objects overlapped substantially (see [Supplemental Data](#)). The auditory objects were presented within virtual auditory space; this was generated by convolving a time-domain representation of the HRTFs [8] for a given location with the waveform for a given object to produce the percept of the sound's emanating from a particular location in extrapersonal space. In experiment 1, the auditory scenes were generated with a standard set of HRTFs (derived from a representative participant) for all participants. The experiment was also conducted on a separate group of participants whose own HRTFs were measured and used to generate the stimuli. All object locations were simulated in the azimuthal plane (range: 0°–350°). No sound was presented at a location directly opposite another sound to exclude any position confusion owing to front-back reversals. Thus, if the sources on the back half-plane were reflected onto the front half-plane, the angular separation between objects ranged from 20° to 180°.

Participants

In experiment 1, 28 participants were tested with standard HRTFs (13 males, 15 females; mean age = 26.3 years). In experiment 2, 12 participants, 8 of whom had participated in experiment 1, were tested (10 males, 2 females; mean age = 31.2) with individualized

HRTFs. In experiment 3, 26 participants were tested (12 males, 14 females; mean age = 26.15), all of whom had participated in experiment 1. All participants were tested audiometrically and had normal hearing thresholds in the frequency range 0.5–8 kHz.

Procedure

All experimental procedures were approved by the Monash University Standing Committee for Ethical Research Involving Humans, and informed consent was obtained from all participants prior to testing. Participants were tested individually. A portable computer (Dell 8100) running DMDX software (J. Forster, University of Arizona) was used to present stimuli and record responses. Auditory stimuli were presented over Sennheiser HD400 headphones (HD520 for individualized HRTF participants). During testing, participants fixated on the center of the computer display. In the directed-attention condition, the name of the cued object appeared at fixation, whereas in the nondirected-attention condition, the phrase “next item” was displayed. Participants were familiarized with the sound library and object names prior to practice and experimental trials. Practice trials included feedback about response accuracy, but experimental trials did not. Participants were allowed to respond at any time during the trial and had no time limit to respond. They were informed about the stimulus conditions prior to each block of trials. Demonstration trials are available online (<http://www.psych.unimelb.edu.au/research/labs/changedeafness.html>).

Measures of performance were the percentage of correct detections of change and d' [11]. Because people use different (more or less conservative) strategies for responding in same-different paradigms, a measure of response criterion (C) was also calculated [11].

Experiment 1

In the nondirected-attention condition, participants made a verbal yes/no judgment after each trial to indicate whether an object had disappeared. In the directed-attention condition, participants made a yes/no judgment to indicate whether the cued object had disappeared. Thus, the task was essentially a same-different task, in which a difference of a specified type had to be detected. The two attention conditions were presented in separate blocks of 60 trials each (45 change, 15 no change), and the order of blocks was alternated across participants. Within each block, there were 20 trials of each scene size (four, six, or eight objects), randomly interspersed.

Experiment 2

In the same-location condition, the position at which all auditory objects were located was varied randomly across trials. The proportion of times a particular position was represented in the same-location condition was equivalent to the proportion of times an object was presented at that position in the different-locations condition. As in experiment 1, the directed- and nondirected-attention conditions were presented in separate blocks, and their order of presentation was varied across participants. The spatial-separation conditions (different locations, same location) were also presented in separate blocks within each of the attention-condition blocks, and their order of presentation was alternated across the two attention conditions. Within each block, there were 20 trials of each scene size (four, six, or eight objects). The different scene sizes were randomly interspersed within each block. Responses were made in the same way as in experiment 1.

Experiment 3

A change in location involved two objects in the initial scene exchanging locations in the second scene. Reciprocal changes in two objects' locations were used (rather than a change in the location of a single object) because changes in the distribution of acoustic energy when one object moved to another location in the second scene could have cued the listener to the change. The range and separation between the two objects involved in a location change were between 40° and 180° in the azimuthal plane. The size of the location change varied randomly within each block. Participants

were informed that each change involved two objects but were only required to detect/identify one of the objects.

Supplemental Data

Supplemental Results and several supplemental figures are available at <http://www.current-biology.com/cgi/content/full/15/12/1108/DC1/>.

Acknowledgments

We thank J. Forster and K. Forster for use of the DMDX software and S. Carlile, C. Chambers, and M. Williams for comments on an earlier version of this manuscript. This work was supported by grants from the National Health and Medical Research Council (Australia) and the Monash University Small Grants Scheme.

Received: January 22, 2005

Revised: May 3, 2005

Accepted: May 3, 2005

Published: June 21, 2005

References

1. Bregman, A.S. (1990). *Auditory Scene Analysis: The Perceptual Organization of Sound* (Cambridge, MA: MIT Press).
2. Darwin, C.J. (1997). Auditory grouping. *Trends Cogn. Sci.* 1, 327–333.
3. Carlyon, R.P. (2004). How the brain separates sounds. *Trends Cogn. Sci.* 8, 465–471.
4. Vitevitch, M.S. (2003). Change deafness: The inability to detect changes between two voices. *J. Exp. Psychol. Hum. Percept. Perform.* 29, 333–342.
5. Desimone, R., and Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annu. Rev. Neurosci.* 18, 193–222.
6. Kastner, S., and Ungerleider, L.G. (2000). Mechanisms of attention. *Annu. Rev. Neurosci.* 23, 315–341.
7. Mack, A., and Rock, I. (1998). *Inattentional Blindness* (Cambridge, MA: MIT Press).
8. Yost, W.A. and Popper, A., eds. (1996). *Virtual Auditory Space: Generation and Applications*. (Austin, TX: R.G. Landes).
9. Wightman, F.L., and Kistler, D.J. (1993). Sound localization. In *Human Psychophysics*, W.A. Yost, A. Popper, and R.R. Fay, eds. (New York: Springer), pp. 155–192.
10. Martin, R.L., McAnally, K.I., and Senova, M.A. (2001). Free-field equivalent localization of virtual audio. *J. Audio Eng. Soc.* 49, 14–22.
11. Macmillan, N.A., and Creelman, C.D. (1991). *Detection Theory: A User's Guide* (Cambridge, UK: Cambridge University Press).
12. Schlauch, R.S., and Hafter, E.R. (1991). Listening bandwidths and frequency uncertainty in pure-tone signal detection. *J. Acoust. Soc. Am.* 90, 1332–1339.
13. Scharf, B. (1998). Auditory attention: The psychoacoustical approach. In *Attention*, H. Pashler, ed. (East Sussex, UK: Psychology Press), pp. 75–117.
14. Mondor, T.A., and Zatorre, R.J. (1995). Shifting and focusing auditory spatial attention. *J. Exp. Psychol. Hum. Percept. Perform.* 21, 387–409.
15. Darwin, C.J., and Hukin, R.W. (1999). Auditory objects of attention: The role of interaural time differences. *J. Exp. Psychol. Hum. Percept. Perform.* 25, 617–629.
16. Huron, D. (1989). Voice denumerability in polyphonic music of homogenous timbres. *Music Perception* 6, 361–382.
17. Yantis, S., and Jonides, J. (1984). Abrupt visual onsets and selective attention: Evidence from visual search. *J. Exp. Psychol. Hum. Percept. Perform.* 10, 601–621.
18. Handel, S.J. (1989). *Listening: An Introduction to the Perception of Auditory Events* (Cambridge, MA: MIT Press).
19. Rensink, R.A., O'Regan, J.K., and Clark, J.J. (1997). To see or not to see: The need for attention to perceive changes in scenes. *Psychol. Sci.* 8, 368–373.
20. O'Regan, J.K., Rensink, R.A., and Clark, J.J. (1999). Change-blindness as a result of “mudsplashes”. *Nature* 398, 34.
21. Simons, D.J., Franconeri, S.L., and Reimer, R.L. (2000). Change blindness in the absence of a visual disruption. *Perception* 29, 1143–1154.
22. Becker, M.W., Pashler, H., and Anstis, S.M. (2000). The role of iconic memory in change-detection tasks. *Perception* 29, 273–286.