

Auditory attention causes visual inattention blindness

Silvia Pizzighello, Paola Bressan

Dipartimento di Psicologia Generale, Università di Padova, Via Venezia 8, 35131 Padova, Italy; e-mail: paola.bressan@unipd.it

Abstract. When engaged in a visual task, we can fail to detect unexpected events that would otherwise be very noticeable. Here we ask whether a common auditory task, such as that of attending to a verbal stream, can also make us blind to the presence of visual objects that we do not anticipate. In two experiments, 120 observers watched a dynamic display while performing either a visual or an auditory attention task, or both simultaneously. When observers were listening to verbal material, in order to either understand it or remember it (auditory task), their probability of detecting an unexpected visual object was no higher than when they were counting bounces of moving items (visual task), although in the former case the observers' eyes and attention could move around the display freely rather than remaining focused on tracked items. Previous research has shown that attending to verbal material does not affect responses to lights flashing at irregular intervals, suggesting that driving performance is not hampered by listening. The lights, however, were expected. Our data imply that listening to the radio while driving, or to a portable audio player while walking or biking, can impair our reactions to objects or events that we do not expect.

1 Introduction

Evidence is accumulating that the focusing of attention on one object or event can render people 'inattentionally blind' to novel unexpected visual stimuli, even when these are quite salient. In the experiments that provided the most striking demonstrations of this phenomenon, participants were engaged in some selective-looking task, such as assessing which arm of a briefly flashed cross was longer (Mack and Rock 1998), or counting the number of either bounces made by moving items (Most et al 2001; 2005) or ball passes between two teams (Simons and Chabris 1999). In these conditions, many participants failed to notice conspicuous unexpected events such as, respectively, a small square appearing next to the flashed cross, a uniquely colored object crossing the screen, and a gorilla intruding on to the sports ground and stopping in the middle to thump its chest.

This peculiar type of blindness shows that our impression that we are constantly representing the world as complete and fully detailed is illusory. To this, one might object that the unanticipated event goes undetected simply because the observers' attention is glued to the objects they are requested to focus on—the flashed cross, the bouncing items, the moving ball. Although the novel item can be missed even when it appears at fixation (eg Koivisto et al 2004; Strayer et al 2003), there are indications that this happens exactly because observers try to concentrate attention on their assigned target rather than on where they are fixating (Mack and Rock 1998). If this were the correct explanation, one would expect a significant reduction

in inattentional blindness if observers were not asked to maintain attention on parts of the scene other than the one where the new object appears; for example, if the main task were non-visual, and the visual display could be explored without constraints. Alternatively, the unexpected event might be missed because the visual attention assigned to the task is unavailable for a new object presented in the same modality (Duncan et al 1997; see also Rees et al 2001). This also leads to the prediction that inattentional blindness would essentially disappear if observers were engaged in a non-visual, as opposed to a visual, attentional task.

A strictly related question is, will concurrent engagement in two tasks, one visual and one non-visual (like listening to the traffic bulletin while driving) augment inattentional blindness (like failing to notice an ice patch lying on the street, say, thirty meters ahead)? And, would this depend on the type of auditory task (eg, on whether we are trying to memorize the steps of an itinerary someone is dictating us over the cell phone, as opposed to simply comprehend what she is telling us)? The existence of cross-modal attentional links between vision and audition has been shown across a wide range of dual-task situations (see Spence and Driver 2004). For example, when people simulate driving and simultaneously carry out a verbal task, both performances are impaired (Horswill and McKenna 1999). However, it is not known whether detection of an *unexpected* event would also be affected. It has been argued that simply attending to verbal material, without active engagement in a conversation, does not interfere with driving (Strayer and Johnston 2001). This conclusion was based on a study in which participants tracked a target that flashed red or green at irregular intervals, and were asked to respond to red by pressing a button. No impairment in reaction to red lights was found when participants simultaneously listened to radio broadcasts, or to a book on tape. However, (a) the critical events were not unexpected, only the time of their occurrence was; and (b) the critical events occurred where the participants' eyes and attention were both focused.

In the work reported here, we searched for an answer to the above questions by using two types of auditory attention task (listening to a few sentences in order to understand them, listening to a list of words in order to remember them) in place of, or in addition to, a visual attention task (counting bounces in a dynamic display).

2 Experiment 1

2.1 Method

2.1.1 Participants

Ninety participants (32 males and 58 females, mean age 37 years) with normal or corrected-to-normal vision were tested individually. They were randomly assigned to one of three conditions: (a) visual, (b) auditory, and (c) visual and auditory (dual task). Within the auditory task conditions, half of the participants were assigned to the comprehension condition and the other half to the recall condition (described in the next section). There were 30 participants per condition.

2.1.2 Stimuli and procedure

The visual stimuli we used were similar to those of Most et al (2001), and were presented on a portable Toshiba Satellite 1800-412 computer with a 14" display. On each of the five experimental trials, four black (luminance=1.0 cd/m²) and four white (luminance=87.4 cd/m²) L and T shapes moved independently on random paths, at variable velocities, against a 10.6x8.0 cm grey (luminance=15.8 cd/m²) background. The third, fourth and fifth trials also contained a light grey (luminance=42.3 cd/m²) cross with the same horizontal and vertical extent as the L's and T's, ie 8 mm, and the same thickness, ie 2 mm. As they moved, the black and white shapes could partially occlude each other, and occasionally bounce off the edges of

the display window.

Auditory stimuli were prepared using the *Italian Assistant, Language Assistant Series* software. They consisted of five short stories in Italian (about 26 words each) and five lists of Italian words (14 words for each list), uttered by a computerized female voice. Each story, or list, lasted 12 seconds. In the auditory task conditions, each of the five trials was coupled with a different story (comprehension condition) or list of words (recall condition). Participants were told they had to listen because their comprehension, or recall, would be tested at the end of the trial.¹

We prepared five separate trials, that were presented in the same order to all participants. The number of bounces was 8 on the first trial, 5 on the second, 6 on the third, 7 on the fourth and fifth. Each trial lasted 12 seconds.

The 90 participants were distributed across five conditions. In the visual-task condition, 30 participants were instructed to watch the display and keep a silent tally, using their fingers, of the number of times that the white letters bounced off the edges of the display window; after each trial, they reported the number of bounces they had seen. In the auditory-task (comprehension) condition, 15 participants watched the display and listened to short stories; after each trial, they answered three questions about the story. In the auditory-task (recall) condition, 15 participants watched the display and listened to lists of words; after each trial, they recalled as many words as possible. In the dual-task (comprehension) condition, 15 participants counted the number of bounces made by the white shapes and listened to short stories; after each trial, they reported the number of bounces they had seen and answered three questions about the story. In the dual-task (recall) condition, 15 participants counted the number of bounces made by the white shapes and listened to lists of words; after each trial, they reported the number of bounces they had seen and recalled as many words as possible.

All observers viewed the display from a distance of about 60 cm and completed five consecutive trials. The first two trials contained no unexpected event. Approximately 2.45 seconds into the third trial (the 'critical trial'), the grey cross unexpectedly entered the display from the right side, traversed the screen horizontally along a virtual midline and exited on the left side (see figure 1). The cross remained visible for 7.15 seconds. After this trial, observers answered a questionnaire adapted from Most et al (2005). They were asked to report whether they had seen anything other than the black and white L's and T's, something that was missing in the first two trials. If the answer was yes, they were asked to describe the color, motion direction, and shape of the object. The shape could be picked from among four different shapes, graphically represented in the questionnaire: an E-shape, a cross, a heart, and a triangle.

Participants then completed a fourth trial (called 'divided-attention trial', because the questionnaire had indirectly alerted them to the possibility that a novel object could appear), after which they answered a second questionnaire, identical to the first. On the fifth and last trial (the 'full-attention' trial), participants were simply asked to view the display, without performing any task. After this trial they answered a final questionnaire, identical to the previous two. Participants were debriefed at the end. Nobody reported to have been familiar with 'inattention blindness' concepts or experiments.

¹ An example of short story would be, "Joe has to go to the drugstore to buy some gauze and plasters. To get there he must turn right at the crossing's traffic-lights". The comprehension questions for this story would be: (1) "Where does this person have to go?"; (2) "What does he have to buy?"; (3) "In which direction does he have to turn to get there?". An example of word list would be, "Flower, Bag, Mountain, Dog, Chair, Pen, Blackboard, Puppy, Book, Window, Candle, Light, Wall, Airplane".

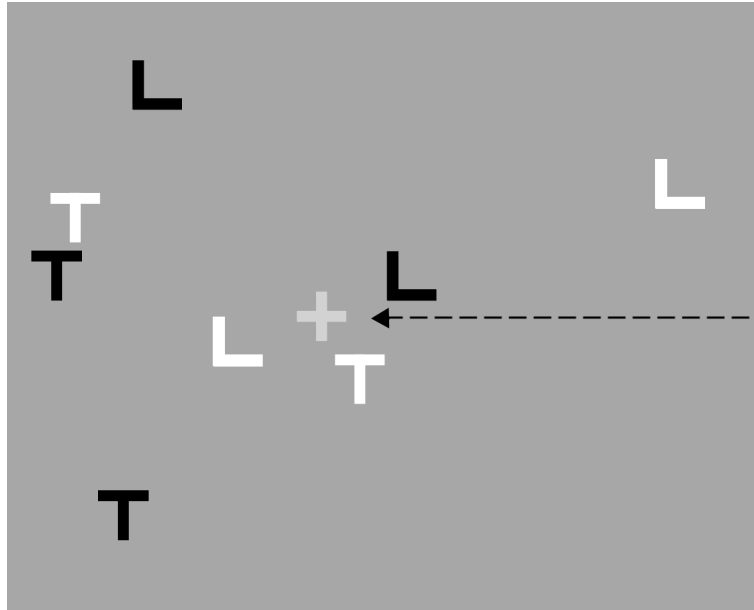


Figure 1. A single frame of one of the five dynamic displays used in the Experiment. The arrow shows the path traveled by the cross, and was not present in the actual display. The white and black shapes moved around on random paths, occasionally touching and bouncing off the edges of the display window.

2.2 Results

Averaging across all conditions, the cross was noticed (ie, the shape was correctly reported) by 38% of the participants on the critical trial, by 72% on the divided-attention trial, and by 100% on the full-attention trial. On the critical trial, only 15 individuals out of 90 correctly described all three attributes (shape, color, and direction of motion) of the unexpected object. Detailed data per trial and experimental condition are shown in Table 1.

condition	critical trial					divided-attention trial			
	seen something new?	shape	color	motion	all three attributes	shape	color	motion	all three attributes
visual	20	17	12	15	8	23	14	19	11
dual	10	8	3	3	1	18	8	16	8
auditory (expt 1)	12	9	7	7	6	24	17	23	16
auditory (expt 2)	22	21	22	17	17	25	24	24	22

Table 1. Number of observers that (i) reported seeing something new on the critical trial, that is gave an affirmative answer to the first question of the questionnaire; (ii) reported correctly the shape, color, and direction of motion of the new object, on the critical trial and on the divided-attention trial. Data are shown separately for the three conditions of Experiment 1 and for the condition of Experiment 2. Each condition included 30 participants.

Performance in the auditory task. Comprehension and recall performances, measured respectively as the number of questions correctly answered and the number of words correctly recalled, were analyzed with two separate repeated-measures ANOVAs, where the within-subjects factor was Cross Presence (pre-critical vs critical trial) and the between-

subjects factors were Condition (auditory-only vs dual task) and Inattentional Blindness (noticers vs nonnoticers). An additional ANOVA was performed on bounce counting accuracy, with a within-subjects factor of Cross Presence (pre-critical vs critical trial) and a between-subjects factors of Condition (visual-only vs dual task) and Inattentional Blindness (noticers vs nonnoticers).

Performance in the auditory task (both comprehension and recall) was reduced when the visual and auditory tasks were combined, $F(1,26)= 9.59$, $p=0.005$ and $F(1,26)= 17.95$, $p<0.0001$ respectively.

The number of words correctly recalled dropped from the pre-critical to the critical trial, $F(1,26)= 5.56$, $p=0.026$. (In the comprehension condition, where the answers to the final questions could partly be inferred from the general context, there was no significant effect.) A marginally significant interaction emerged between Cross Presence and Condition, $F(1,26)=4.02$, $p=.055$, due to the fact that recall worsened in the auditory-only condition (from 5.07 to 3.87 words) but not in the dual-task one (2.73 words in both cases). In all likelihood the lack of a performance drop in the latter case reflects a floor effect, due to the high attentional demands of the dual task.

The decline in recall was essentially due to noticers, as shown by a marginally significant interaction between Inattentional Blindness and Cross Presence, $F(1,26)= 3.13$, $p= 0.09$. More specifically, noticers recalled fewer words in the critical relative to the pre-critical trial (the difference amounted to 1.75 words and was significant, $p= 0.04$), whereas the performance of nonnoticers did not change (the difference amounted to 0.18 words and was nonsignificant, $p>.1$). This might be simply explained by the fact that, unlike nonnoticers, noticers either became temporarily distracted from the auditory task when noticing the cross (interference), or kept a trace of the cross in memory until the end of the trial, in addition to the traces of the words (higher memory load).

Performance in the visual task. Performance in the visual task was also reduced when the visual and auditory tasks were combined, with an average of 1 error (mean absolute deviation from correct number of bounces) in the dual task vs 0.5 errors in the visual-only one, $F(1,56)=5.12$, $p= .028$. On the critical trial, participants who did not notice the cross made significantly *more* counting errors than noticers (in both the visual and dual conditions), suggesting that the unexpected object might have been subliminally monitored, thereby consuming attentional resources. This finding, together with supporting results coming from additional experiments, has been reported elsewhere (Bressan and Pizzighello, in press) and will not be expanded upon here.

Inattentional blindness. From the standpoint of their effects on inattentional blindness, the comprehension and recall conditions were not significantly different, either alone, $\chi^2<1$, or within the dual task, $\chi^2<1$; hence, they were combined. Unsurprisingly, observers were much less likely to notice the cross when they had to simultaneously attend to two tasks, one visual and one auditory, rather than to a visual task only, $\chi^2(1, N=60)=5.55$, $p=0.018$. (See Table 1.) Surprisingly, however, there was no significant difference in the amount of inattentional blindness between people engaged in the dual task and people engaged in the auditory task alone, $\chi^2<1$. This was true only as long as the stimulus was unanticipated: in the divided-attention trial, the gap between the two conditions disappeared, $\chi^2<1$, and there was a tendency for the cross to be noticed less often when the task was dual as opposed to auditory, $\chi^2(1, N=60)=2.86$, $p=0.09$.

The dual task also hampered detailed perception of the cross: the number of perfect noticers (participants who correctly reported all three attributes of the cross) was significantly smaller in the dual-task condition than in either the auditory, $\chi^2(1, N=60)=4.04$, $p=.044$, or the

visual one, $\chi^2(1, N=60)=6.40, p=.011$.

2.3 Discussion

Only one-third of our observers reported noticing something unexpected in the conditions in which they were listening to verbal material. Of course they might have perceived the cross, but then forgotten about it by the time they were questioned (ie, inattention blindness may actually be “inattention amnesia”, Wolfe 1999). However, our finding that bounce-counting accuracy in the critical trial deteriorated in nonnoticers, but not in noticers, is inconsistent with the idea that the unattended item was perceived but then forgotten (see also Bressan and Pizzighello, in press). This idea presupposes that all observers are actually noticers. Yet if this were the case, performance should, if anything, be worse for noticers (who keep a trace of the object in memory until the end) than for nonnoticers (who dispose of it).

We found that inattention blindness was as likely in the dual as in the auditory-only condition: adding an auditory task to a visual task worsened inattention blindness, but adding a visual task to an auditory one did not. A possible explanation is that of a ceiling effect. It has been shown that perception of irrelevant distractors is eliminated under conditions of high attentional load in an unrelated task (Rees et al 2001). The load of the auditory task might be so high that it engages attention fully, exhausting available capacity and making the supplemental visual task redundant. However, the higher attentional load of the dual task relative to the auditory-only task was clearly revealed by a corresponding decrease in both word retention and probability of full perception (or retention) of the cross. We must conclude that the auditory task, in itself, consumed attention only partly, and that attentional capacity was not at ceiling.

Folk et al (1992; 1993; 2002) argued that, in order to capture attention, a target must be part of a top-down attentional set. When the task is visual and the distractor is auditory, they are less likely to be part of the same attentional set than when both are visual. Since in the auditory-only task the visual unexpected object is not part of the attentional set, it is also possible that, despite instructions to watch the screen, some participants may have ‘defocused’ the display (by converging their eyes either behind or in front of it) in order to ‘focus’ on the auditory task itself. Whereas participants were instructed to watch the screen and did keep their eyes on the monitor, by the third (critical) trial they would have realized that the visual stimulus was irrelevant to the task.

For this reason we ran a new experiment, where participants were explicitly told that it was crucial that they watched the videoclip from start to finish, and that they would be tested with a simple question about it at the end of each trial.

3 Experiment 2

3.1 Method

3.1.1 Participants

Thirty participants (1 male and 29 females, mean age 22 years) with normal or corrected-to-normal vision were tested individually. Five additional participants (3 noticers and 2 nonnoticers) were replaced because, when questioned at the end, they admitted they had not watched all videoclips from start to finish. All participants were tested with the auditory task: 15 participants were randomly assigned to the comprehension condition and 15 to the recall condition.

3.1.2 Stimuli and procedure

Apparatus, stimuli, and procedure were identical to those used in the auditory-only condition of Experiment 1, with the following difference. In Experiment 1 participants were asked to watch the screen, but no attempt was made to check to what extent they had actually done so.

In Experiment 2 participants were explicitly instructed to watch (without counting or performing any specific visual task) each videoclip for its whole duration. They were told that (a) this was important, and (b) they would be tested with a question about it at the end of each trial.

As in Experiment 1, after each trial participants answered three questions about the story (comprehension condition) or recalled as many words as possible (recall condition). In addition, after the first (practice) trial they were asked which color the letters were, and after the second (noncritical) trial they were asked which color the background was. After the third (critical), the fourth (divided-attention), and the fifth (full-attention) trials, participants answered the same questionnaire as in the previous experiment.

At the end of the experiment participants were asked if they had indeed watched the videoclip from start to finish, and fully debriefed. Nobody reported to have been familiar with 'inattentional blindness' concepts or experiments.

3.2 Results

The number of participants that showed inattentional blindness in the trials of interest (critical and divided-attention; all subjects reported the cross in the full-attention trial) is given in Table 1. In the critical trial, this number was lower than in either the dual-task or auditory-only conditions of Experiment 1 ($\chi^2(1, N=60)=11.28, p=.001$, and $\chi^2(1, N=60)=9.6, p=.002$ respectively). Interestingly, although all 30 participants reported that they had watched each videoclip from start to finish, only 24 of them correctly reported the colors of both sets of moving letters, ie black and white (a further 4 participants recalled one color only), and only 21 correctly reported the color of the background, ie grey. On the other hand, inattentional blindness was not linked to a scarce attentiveness to the visual display in general: the probability of missing the cross was basically the same ($\chi^2<1$) among those who responded correctly to both control questions (17 participants, 5 nonnoticers) and those who did not (13 participants, 3 nonnoticers).

Like in Experiment 1, the decrease in number of correctly recalled words from the pre-critical (4.53) to the critical (2.00) trial was significant, $F(1,13)= 16.61, p=.001$. Because participants were forced to watch the display for its whole duration, we expected a worse recall performance in Experiment 2 than in Experiment 1. Across trials, participants did indeed recall fewer words in Experiment 2 (3.33 on average) than in Experiment 1 (4.25 on average), $F(1,28)= 12.42, p=.001$.

3.3 Discussion

Experiment 2 confirmed the finding of Experiment 1 that the engagement of auditory attention, per se, can block an unexpected visual object from reaching awareness, even when no specific visual task is simultaneously carried on. The comparison between the two experiments suggests that, in order to focus on a verbal stream, people instinctively disengage visual attention—even when they are explicitly asked not to, and to the extent that they can get away with it (hence, less in Experiment 2 than in Experiment 1). This automatic disengagement of visual attention pays back: the number of words correctly recalled was indeed significantly smaller in Experiment 2, that is when participants were forced to watch the visual event from start to finish. Nonetheless, a sizeable amount of inattentional blindness occurred even in this case, although participants not only were told that watching the display was a crucial part of the experimental task, but also reported, at the end of the experiment, to have fully followed the instructions.

7 Conclusions

Half of our observers failed to notice the unexpected visual object in the auditory-only

condition, even though, in the absence of a tracking task, they were free to move both their visual attention and their eyes over the display. This shows that inattention blindness does not need allocation of focused visual attention to concurrent stimuli (on a 'add here/subtract there' same-modality principle, as generally implied in the literature), but can be induced by allocation of auditory attention as well. Actually, the engagement of auditory attention was no less effective than the engagement of visual attention. Our experiments indicate that (a) in order to pay attention to a verbal stream observers automatically withdraw attention from the visual scene; and (b) they do this even when they are explicitly required to watch the scene, and are convinced to have done so.

Situations in which we must pay attention to visual and verbal stimuli simultaneously are remarkably common outside laboratories. The importance of understanding how such tasks interfere with one another and with our ability to detect new visual objects is especially obvious when the consequences can be tragic, as in the case of driving accidents. Use of a cellular phone is associated with a fourfold increase in the likelihood of a crash that will result in hospital attendance (McEvoy et al 2005). The hazard is totally independent of the type of telephone—hand-held and hands-free devices entail equivalent risks. More worryingly, a conversation with a passenger has about the same effects as a conversation on the cell phone; and even information-processing tasks much less engaging and emotionally loaded than a conversation, such as mental arithmetic or word games, entail significant costs (see Horrey and Wickens 2006 for a meta-analysis). These costs are manifested primarily in terms of longer reaction times.

Nonetheless, no impairment in reaction times to irregularly occurring red lights was found for people who performed a visual tracking task and simultaneously listened to radio broadcasts, or to books on tape (Strayer and Johnston 2001). The authors concluded that simply attending to verbal material, without being involved in a conversation, does not affect driving performance. However reassuring it may appear to those of us who drive, this conclusion needs to be qualified by adding that this lack of interference was found for stimuli whose occurrence was erratic but expected. Our work crucially complements these data by indicating that listening to the radio while driving, or to a portable audio player while walking or biking, may impair our reactions to stimuli that we do not expect.

Acknowledgments. Part of this research was submitted as an undergraduate thesis by Silvia Pizzighello to the Department of General Psychology at the University of Padova, under the supervision of Paola Bressan. We are indebted to Peter Kramer, Stefano Vezzani, Steven Most, James Cutting, and Carlo Umiltà for discussion; we also thank Charles Folk and Michael Silverman for their thoughtful comments.

References

- Bressan P, Pizzighello S, "The attentional cost of inattention blindness" *Cognition* (forthcoming) doi:10.1016/j.cognition.2007.03.001
- Duncan J, Martens S, Ward R, 1997 "Restricted attentional capacity within but not between sensory modalities" *Nature* **387** 808-810
- Folk C L, Leber A, Egeth H, 2002 "Made you blink! Contingent attentional capture produces a spatial blink" *Perception & Psychophysics* **64** 741-753
- Folk C L, Remington R W, Johnston J C, 1992 "Involuntary covert orienting is contingent on attentional control settings" *Journal of Experimental Psychology: Human Perception and Performance* **18** 1030-1044
- Folk C L, Remington R W, Johnston J C, 1993 "Contingent attentional capture: A reply to Yantis" *Journal of Experimental Psychology: Human Perception and Performance* **19** 682-685

- Horrey W J, Wickens C D, 2006 "Examining the impact of cell phone conversations on driving using meta-analytic techniques" *Human Factor* **48** 196-205
- Horswill M S, McKenna F P, 1999 "The effect of interference on dynamic risk-taking judgments" *British Journal of Psychology* **90** 189-199
- Koivisto M, Hyönä J, Revonsuo A, 2004 "The effects of eye movements, spatial attention, and stimulus features on inattentive blindness" *Vision Research* **44** 3211-3221
- Mack A, Rock I, 1998 *Inattentive Blindness* (Cambridge, MA: The MIT Press)
- McEvoy S P, Stevenson M R, McCartt A T, Woodward M, Haworth C, Palamara P., Cercarelli R, 2005 "Role of mobile phones in motor vehicle crashes resulting in hospital attendance: a case-crossover study" *British Medical Journal* **331** 428-432
- Most S B, Scholl B J, Clifford E R, Simons D J, 2005 "What you see is what you set: Sustained inattentive blindness and the capture of awareness" *Psychological Review* **112** 217-242
- Most S B, Simons D J, Scholl B J, Jimenez R, Clifford E, Chabris C F, 2001 "How not to be seen: The contribution of similarity and selective ignoring to sustained inattentive blindness" *Psychological Science* **12** 9-17
- Rees G, Frith C, Lavie N, 2001 "Processing of irrelevant visual motion during performance of an auditory attention task" *Neuropsychologia* **39** 937-949
- Simons D J, Chabris C F, 1999 "Gorillas in our midst: Sustained inattentive blindness for dynamic events" *Perception* **28** 1059-1074
- Spence C, Driver J, 2004 *Crossmodal space and crossmodal attention* (Oxford, UK: Oxford University Press)
- Strayer D L, Drews F A, Johnston W A, 2003 "Cell phone-induced failures of visual attention during simulated driving" *Journal of Experimental Psychology: Applied* **9** 23-32
- Strayer D L, Johnston W A, 2001 "Driven to distraction: Dual-task studies of simulated driving and conversing on a cellular phone" *Psychological Science* **12** 462-466
- Wolfe, J. M. (1999). Inattentive amnesia. In V. Coltheart (Ed.), *Fleeting Memories: Cognition of Brief Visual Stimuli* (pp. 71-94). Cambridge, MA: MIT Press.