Restricted attentional capacity within but not between sensory modalities

John Duncan*, **Sander Martens**† & **Robert Ward**‡

* MRC Applied Psychology Unit, 15 Chaucer Road, Cambridge CB2 2EF, UK
† Unit of Experimental and Theoretical Psychology, Faculty of Social Sciences, Leiden University, PO Box 9555, 2300 RB Leiden, The Netherlands
‡ School of Psychology, University of Wales, Bangor, Gwynedd LL57 2DG, UK

Restrictions to attentional capacity are revealed by the interference that commonly results when two sensory inputs must be identified at the same time[1]. To investigate this phenomenon within and between modalities, we presented streams of visual and/or auditory inputs, containing occasional targets to be identified and recalled. For two visual or two auditory streams, identification of one target produced a sustained reduction in the ability to identify a second, the period of interference lasting for several hundred milliseconds. Subjectively, when attention was assigned to one target it was temporarily unavailable for another. In contrast, there was no such time-locked interference between targets in different modalities. The results suggest a modality-specific restriction to concurrent attention and awareness; visual attention to one simple target does not restrict concurrent auditory attention to another.

Rather generalized limitations on concurrent attention and awareness have often been proposed[2]. According to such views, attention to any target stimulus or event should interfere with concurrent attention to any other subsequent target, regardless of the targets' modalities. Very few studies, however, have compared interference within and between modalities, leaving the question unresolved. As one suggestive exception, Treisman and Davies[3] found that searching for a target across two inputs in different modalities was more efficient than searching across inputs in the same modality. Here we extend this early study with exact measurements of the time-course of interference produced by one attended target on another, as a function of whether targets are presented within or between modalities. Interference between concurrent tasks can occur for many reasons[4]; measures of interference time-locked to a specific sensory input are useful in localizing the specific attentional demands of its processing[4].

Figure 2 Example trial for single-modality visual experiment. Throughout each trial, the display contained a central fixation dot and four four-dot markers indicating stimulus positions. Overall display dimensions were ~4.3° square. As in the auditory experiment, there were two streams of input. Each stream was a series of 'frames', each presented for 150 ms and separated by a blank interval of 100 ms from the next. For the horizontal stream, each frame contained two letter strings, one in the left position and one in the right. In non-target frames, each string was a row of three xs, the whole string approximately 0.6° in length. In target frames, at random either the left or right non-target was replaced by the target word 'nap' or 'nab'. For the vertical stream, strings were presented in upper and lower positions, and the target was the word 'cot' or 'cod'. As indicated, frames for horizontal and vertical streams were partly overlapping in time. Thus, for part of the duration of any given frame, the other two display locations were empty (as illustrated), whereas for the remainder of the duration those locations also contained letter strings. Other details of timing were as in the auditory experiment, except that, to keep performance to an appropriate level, target words were presented for only 120 ms, and followed after this SOA by a 50 ms non-target (three xs) in the same location. SOA between the target and the next whole non-target frame, however, was preserved at the usual 250 ms.
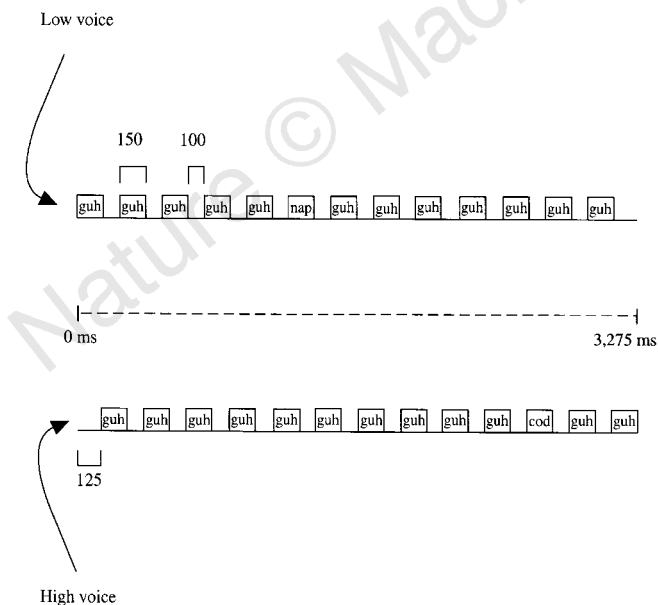

Figure 1 Example trial for single-modality auditory experiment. Two streams of speech were presented concurrently, one in a low and one in a high voice. Each stream consisted of identical repetitions of the syllable 'guh' (non-targets), with a single target word ('nap' or 'nab' for the low voice, 'cot' or 'cod' for the high voice) embedded within it. Syllables and words lasted for 150 ms each, and were separated by silent intervals of 100 ms. One stream, chosen at random on each trial, began 125 ms before the other, breaking synchrony of the two. The first stream to begin also contained the first target, presented after five non-targets. The second target was presented in the other stream, following delays of 125, 375, 625 or 1,375 (illustrated) ms measured from first to second target onset (stimulus onset asynchrony or SOA). After the second target, each stream was completed by two final non-targets.
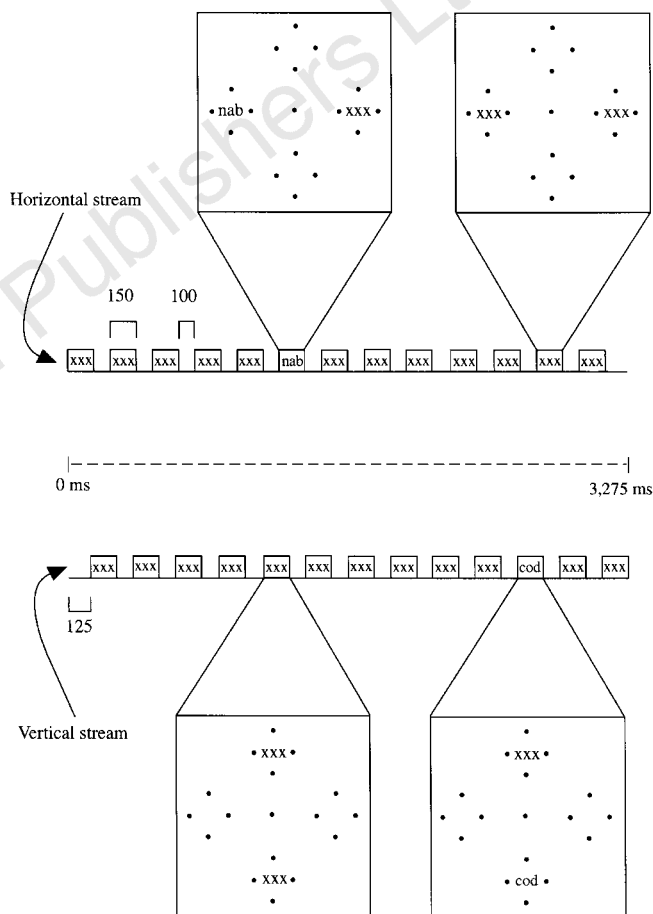
Our methods followed those recently used to measure the time course or dwell time of visual attention[5-7]. Single-modality cases are illustrated in Figs 1 and 2. In the auditory case (Fig. 1), concurrent input streams were spoken by high and low voices. Each stream consisted of a string of non-targets (the spoken syllable 'guh'), with a single target word embedded somewhere within it. In the crucial divided-attention condition, the subject's task was to identify both targets; we measured the accuracy of identifying each target and, by varying their temporal separation, the time course of their mutual interference, or the dwell time of auditory attention. Focused-attention conditions, in which one stream of input was disregarded and only the target in the relevant stream was to be reported, were used as controls. The single-modality visual case (Fig. 2) was similar. Each stream of input was a series of briefly flashed letter strings presented one after the other. To ensure that central fixation was always maintained, the elements of each stream were not single strings but pairs of strings, either to left and right of fixation (horizontal stream) or above and below (vertical stream). Non-targets were rows of three xs, and targets again were words. Again the task in the crucial divided-attention condition was to identify both targets, one from the horizontal and one from the vertical stream.
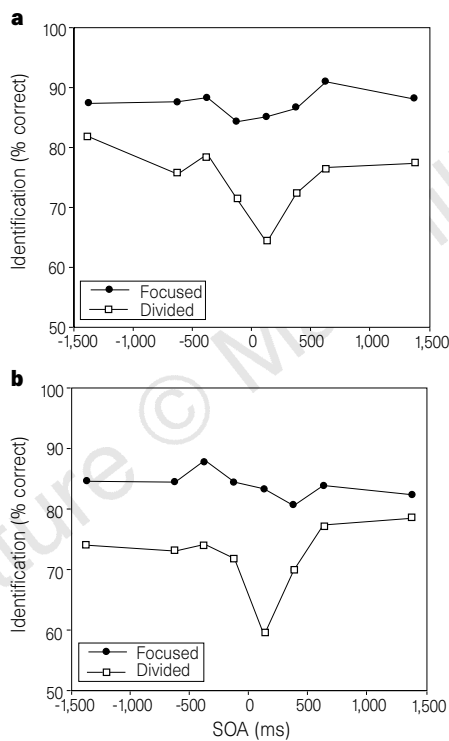
Figure 3 shows results from these single-modality experiments, comparing divided-attention conditions with focused-attention controls. Results in the two modalities were closely similar. Control conditions were more accurate overall, reflecting generalized dual-task decrements[4,8,9] not time-locked to target onsets. Of primary interest is the time course of interference between one target and another in the divided-attention case. In either modality, accuracy was substantially reduced when a first target was followed within a few hundred milliseconds by a second target in the other stream. The results reflect the dwell time of attention on a first target and the subsequent release of attentional capacity for the next. Such interference was entirely absent from control conditions.

Figure 4 shows results from a mixed-modality experiment, in which one stream was auditory and the other was visual. Again there was a generalized dual-task decrement in the divided-attention condition. This time, however, there was absolutely no interference time-locked to target presentation. Attending to a target in one modality left concurrent identification of a second target in a different modality undisturbed.

We can consider the relation of these results to a variety of other phenomena in the dual-task literature. Previous findings of small overall differences between focused and divided attention, even
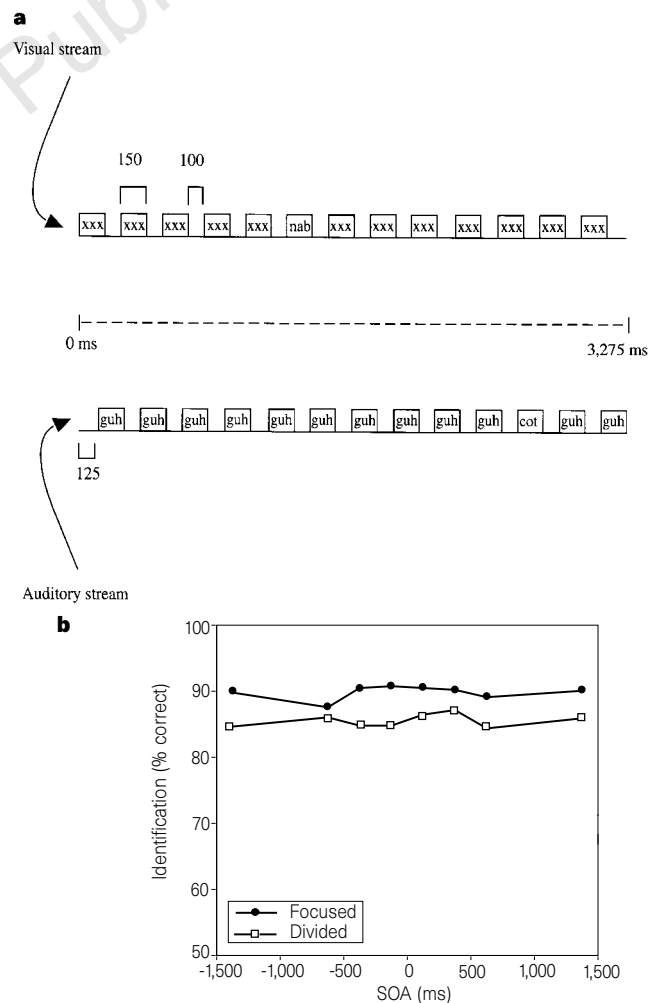
Figure 3 Mean accuracy (percentage correct word identification) in single-modality experiments. Negative SOAs refer to the first target in each trial; positive SOAs refer to the second target. For each experiment, the two streams (low and high voice for the auditory experiment, vertical and horizontal for the visual experiment) were scored independently; values shown are means across the two streams. **a**, Auditory experiment. Both the main effect of condition (focused compared with divided attention) and its interaction with SOA were significant (F-test, $P < 0.002$ in each case). In the divided-attention condition, accuracy at each SOA was compared with asymptote, defined as mean accuracy at SOAs of $\pm 1,375$ ms. Significant reductions occurred at SOAs of -125, $+125$ and $+375$ ms, $P < 0.01$ or better. A final test showed a significant difference between focused- and divided-attention conditions even at asymptote, $P < 0.01$. **b**, Visual experiment. Again, the main effect of condition and its interaction with SOA were both significant, $P < 0.001$ in each case. In the divided-attention condition, significant reductions compared with asymptote occurred at SOAs of $+125$ ($P < 0.001$) and (marginally) $+375$ ($P < 0.03$) ms. Again, focused and divided attention differed even at the asymptote, $P < 0.05$.

Figure 4 Mixed-modality experiment. **a**, Stimulus streams for a typical trial. The horizontal stream of the visual experiment was combined with the high voice stream of the auditory experiment. Timing details were as before. **b**, Mean accuracy. There was a significant main effect of condition (focused compared with divided attention), $P = 0.02$, but no main effect of SOA or condition–SOA interaction, each $F < 1$. Values are means for auditory and visual streams; except that the auditory task carried the bulk of the dual-task cost, auditory and visual data were similar.

when inputs are in different modalities[3,10,11], are consistent with our own generalized dual-task decrement (Fig. 4). As our results show, it is important that we measured interference time-locked to target presentation. Our experiments deal only with the case of independent visual and auditory inputs; a separate question concerns cross-modal integration when the same event gives rise to input in different modalities[12]. Further sources of dual-task interference might also be important in more complex tasks, or when speeded responses are required[4]. Under the simple conditions of our experiments, however, attention to concurrent targets shows no cross-modal restriction. Neurophysiologically, the results suggest that a major source of attentional restriction must lie in modality-specific sensory systems. □

Methods

In each experiment, concurrent input streams were presented at a rate of 250 ms per item. A single target occurred in each stream. In focused-attention conditions, the task was to identify just the target in one specified stream. In divided-attention conditions, both targets were to be identified.

Streams for the single-modality auditory experiment are illustrated in Fig. 1. To create these sequences, single instances of each target and non-target were recorded, then cut to a length of 150 ms and combined into appropriate streams with Macromedia SoundEdit Pro. These streams were presented over headphones at a loudness of approximately 78 dB sound pressure level, with an Apple Macintosh IIci running Cedrus SuperLab software.

Streams for the single-modality visual experiment are illustrated in Fig. 2. Stimuli in this experiment were presented on an Apple Power Macintosh 8500 running Psyscope software[13]. Exact measurements of display timing revealed small variations around the values specified in the program; for example, when stimulus onset asynchrony (SOA) between the two targets was specified as 125 ms, measured values ranged from 119 to 148 ms, and similar levels of variability were seen in target durations and in SOA between one frame and the next within each stream. Such fluctuations were the same under divided- and focused-attention conditions and can accordingly be ignored.

Streams for the mixed-modality experiment are illustrated in Fig. 4. The experiment was run on an Apple Power Macintosh 8500 running Psyscope software, with temporal variability comparable to the visual experiment.

In all experiments, the subject initiated a trial by pressing the space bar of the computer keyboard. Stimulus streams began after a fixed delay of 250 ms; identification responses were typed in after the streams finished, using keys appropriately labelled for the targets 'cot', 'cod', 'nap' and 'nab'. Under focused-attention conditions, there was a single response identifying the attended target (two-alternative forced choice). Under divided-attention conditions, two responses were typed in, in either order. Subjects were strongly encouraged to take time over their responses, ensuring that typing errors were not made. For comparability, the fixation display (central dot and location markers, Fig. 2) was presented in all three experiments, and central fixation was required throughout the trial.

Each subject served in two focused-attention conditions, one for each stream, and the divided-attention condition. Each experiment began with three practice blocks of 32 trials, one per condition. In the auditory experiment there followed a first set of three experimental blocks (96 trials each), one per condition, followed by a second, similar set; other experiments had only one set of three experimental blocks (128 trials each), one per condition. The order of conditions was fixed for any one subject but counterbalanced across subjects. The experiment was conducted in a single session lasting ~1.5 h.

Single-modality auditory and visual experiments had 12 paid subjects each, aged 19–50 years. The mixed-modality experiment had 18 subjects, aged 19–49 years.

1. Broadbent, D. E. *Perception and Communication* (Pergamon, London, 1958).
2. Posner, M. I. & Petersen, S. E. The attention system of the human brain. *Annu. Rev. Neurosci.* **13**, 25–42 (1990).
3. Treisman, A. M. & Davies, A. in *Attention and Performance IV* (ed. Kornblum, S.) 101–117 (Academic, London, 1973).
4. Pashler, H. Dual-task interference in simple tasks: data and theory. *Psychol. Bull.* **116**, 220–244 (1994).
5. Duncan, J., Ward, R. & Shapiro, K. Direct measurement of attentional dwell time in human vision. *Nature* **369**, 313–315 (1994).
6. Broadbent, D. E. & Broadbent, M. H. P. From detection to identification: response to multiple targets in rapid serial visual presentation. *Percept. Psychophys.* **42**, 105–113 (1987).
7. Raymond, J. E., Shapiro, K. L. & Arnell, K. M. Temporary suppression of visual processing in an RSVP task: an attentional blink? *J. Exp. Psychol: Hum. Percept. Perf.* **18**, 849–860 (1992).
8. Logan, G. D. Attention in character-classification tasks: evidence for the automaticity of component stages. *J. Exp. Psychol: Gen.* **107**, 32–63 (1978).
9. Bourke, P. A., Duncan, J. & Nimmo-Smith, I. A general factor involved in dual task performance decrement. *Q. J. Exp. Psychol.* **49A**, 525–545 (1996).
10. Lindsay, P. H., Taylor, M. M. & Forbes, S. M. Attention and multi-dimensional discrimination. *Percept. Psychophys.* **4**, 113–117 (1968).
11. Massaro, D. W. & Warner, D. S. Dividing attention between auditory and visual perception. *Percept. Psychophys.* **21**, 569–574 (1977).
12. Driver, J. Enhancement of selective listening by illusory mislocation of speech sounds due to lip-reading. *Nature* **381**, 66–68 (1996).
13. Cohen, J. D., MacWhinney, B., Flatt, M. & Provost, J. PsyScope: a new graphic interactive environment for designing psychology experiments. *Behav. Res. Meth. Inst. Comp.* **25**, 257–271 (1993).

The small GTP-binding protein Rab3A regulates a late step in synaptic vesicle fusion

Martin Geppert*, **Yukiko Goda**†, **Charles F. Stevens**† **& Thomas C. Südhof**‡

* *Max Planck Institute for Experimental Medicine, Hermann-Rein-Strasse 3, 37075 Gottingen, Germany*
† *Molecular Neurobiology Laboratory and Howard Hughes Medical Institute, The Salk Institute, La Jolla, California 92037, USA*
‡ *Department of Molecular Genetics and Howard Hughes Medical Institute, The University of Texas Southwestern Medical Center, Dallas, Texas 75235, USA*

The Rab family of low-molecular-mass GTP-binding proteins are thought to guide membrane fusion between a transport vesicle and the target membrane, and to determine the specificity of docking[1-3]. The docking and fusion of vesicles is, however, a complex multistep reaction, and the precise point at which Rab proteins act in these sequential processes is unknown. In brain, the Rab protein Rab3A is specific to synaptic vesicles, whose exocytosis can be monitored with submillisecond resolution by following synaptic transmission. We have now determined the precise point at which Rab3A acts in the sequence of synaptic vesicle docking and fusion by using electrophysiological analysis of neurotransmitter release in Rab3A-deficient mice. Unexpectedly, the size of the readily releasable pool of vesicles is normal, whereas Ca^{2+}-triggered fusion is altered in the absence of Rab3A in that a more-than-usual number of exocytic events occur within a brief time after arrival of the nerve impulse.

We first investigated whether the *rab3A* deletion modifies the properties of the fundamental unit of transmitter release: the quantity of transmitter contained in each vesicle and the basal rate of spontaneous vesicle fusion. We used spontaneous miniature excitatory postsynaptic currents (mEPSCs), each one of which corresponds to exocytosis by an individual vesicle[4] (the neurotransmitter quantum), to compare the amplitudes of individual quantal responses between wild-type and *rab3A* mutant synapses (Fig. 1a). Both the amplitude of mEPSCs and their basal spontaneous rate are unchanged at the *rab3A*-mutant synapses (Fig. 1b).

With rapid repeated use, the average quantity of neurotransmitter released by a synapse declines, a decline that probably represents the depletion of fusion-competent vesicles in the 'readily releasable pool' at the presynaptic terminal[5]. This depletion is faster in *rab3A*-mutant neurons, indicating some alteration in the efficiency of synaptic vesicle trafficking[6]. To determine whether the depletion is faster because the readily releasable pool is smaller than usual, we compared the pool size at synapses formed between cultured hippocampal neurons prepared from wild-type or *rab3A*-mutant mice.