Spatiotemporal attentional constraints in perception

Jeroen S. Benjamins

Universiteit Utrecht

Helmholtz Instituut

Spatiotemporal attentional constraints in perception

Spatieel-temporele beperkingen van aandacht in perceptie

(Met een samenvatting in het Nederlands)

Proefschrift

ter verkrijging van de graad van doctor aan de Universiteit Utrecht
op gezag van de Rector Magnificus, Prof. Dr. J.C. Stoof, ingevolge het
besluit van het College voor Promoties in het openbaar te verdedigen

op 10 januari 2008 des ochtends te 10:30 uur

door

Jeroen Sander Benjamins

geboren op 21 maart 1980, te Groningen

Promotor:	Prof. dr. Frans A.J. Verstraten *Universiteit Utrecht* *Helmholtz Instituut*
Copromotoren:	Dr. Ignace T.C. Hooge *Universiteit Utrecht* *Helmholtz Instituut* Dr. Maarten J. van der Smagt *Universiteit Utrecht* *Helmholtz Instituut*

Het in dit proefschrift beschreven onderzoek werd uitgevoerd met financiële steun van de Nederlandse organisatie voor Wetenschappelijk Onderzoek (NWO).

Printed by Gildeprint, Enschede, the Netherlands

ISBN: 978-90-9022644-6

voor mijn vader en moeder

Table of contents

Chapter 1

Introduction

Introduction

Consider the following situation: you are in your car, waiting for the traffic light to turn green at a busy intersection. Three of your friends are engaged in a lively conversation trying to involve you as well, while you try not to be distracted by flashing advertisement signs along the road. In other words, you get all sorts of sensory input, while you are supposed to do a specific task, which is to get your car from A to B without causing damage. Since you can not process all this information simultaneously (see Wolfe, 1994), you need to make a selection of all input signals and discard the rest. The selected information can then be used to perform the task you need to perform. Put differently, you need to focus on some of the sensory input, while ignoring other input.

This selection of objects and events is determined not purely on whatever salient objects or events stand out the strongest in the current flow of sensory input. Current goals, interests etcetera steer the perceptual system in a top-down manner by directing the senses to a particular part of the world as well. For example, when driving a car the focus tends to be on the redness of the traffic light and not on the redness of advertisement signs. Of course, a loud explosion will definitely direct the senses away from a traffic light and towards whatever exploded. However, the more interesting instances are where much less extravagant events and objects can disturb the apparent top-down control normally exerted over our perception. An example of such instances is the attentional blink (Raymond, Shapiro & Arnell, 1992; Duncan, Ward & Shapiro, 1994), where observers fail to notice a second target that is presented shortly after a first target in stream of rapidly presented sequence of objects. Likewise, research on attentional capture has shown that while doing a certain task on which observers are currently focussing, some irrelevant events and objects can distract from that task, even while observers know that these events are irrelevant (Yantis & Hillstrom, 1994; Yantis & Jonides, 1984).

As the name of the phenomena mentioned above suggest, constraints in perceptual selection might not only be due to constraints in the sensory systems - these have their own limits (Kelly, 1979) - but due to top-down attentional selection mechanisms as well. The interesting question then is how these attentional selection mechanisms work and what their limits are. This thesis focuses on how well perception can be voluntary (top-down) controlled by examining the spatiotemporal limits of attentional selection. To study these limits, we need models of attentional selection. Therefore a short overview of some of these models follows below. After the description of these models, a short overview of temporal and spatial limits of attentional selection will further form the starting point of this thesis.

Models of attention

In the last two to three decades, a number of models of attention (Treisman & Gelade, 1980; Wolfe & Cave, 1989; Wolfe, 1994; Wolfe & Gancarz, 1996) have been proposed to explain how attention operates when selecting events and objects from the incoming sensory input. A well-known and often cited model of attentional selection is the feature integration theory by Anne Treisman (1980). This theory or model states that to be able to select the right objects or events to perform a task on, one has to find the right combination, or as she calls it *conjunction*, of features. Different features are individually processed and are registered separately in feature maps and in a master map of locations. When an object or event is not defined by a single unique feature, the perceptual selection of that object or event can not be retrieved by accessing one feature map. The master map of locations in combination with several feature maps has to be checked for the right conjunction of features. In this model, checking the master map of locations and several feature maps is a top-down attentive process (see Figure 1.1), while perceiving an object or event based on single feature is a preattentive process according to this model.

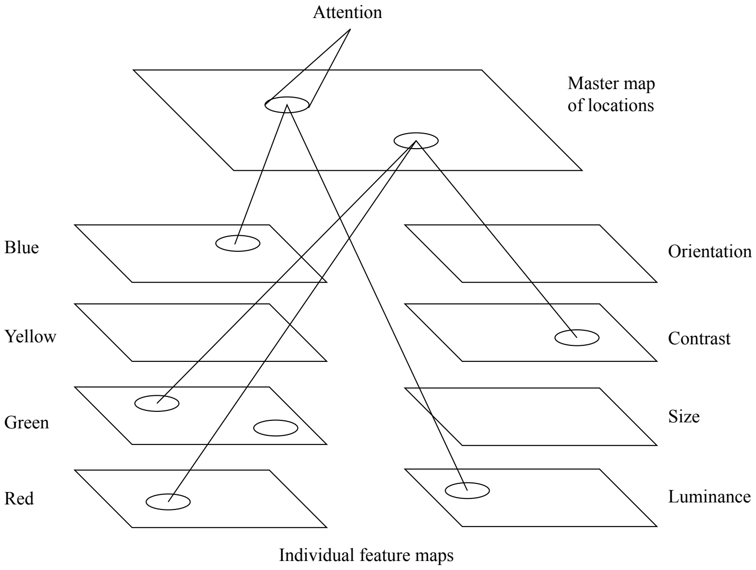

Figure 1.1. Schematic view of Treisman's original feature integration model (1980). Attention acts on the master map of locations to find conjunction of features. If an object is defined by one feature it will be perceived without attention.

Treisman first set out to determine which features were processed preattentively, in other words she determined which features had their own feature map. However, she later attenuated the proposed dichotomy between features being processed either in parallel (or preattentively) or serially (or attentively). She showed that parallel visual search for a single feature-defined target can become serial (Treisman & Gormican, 1988) by varying target non-target similarity.

Strict serial processing (cf. Ward & McLelland, 1989) has not been found in many empirical studies and more and more, the trend is to view serial and parallel processing

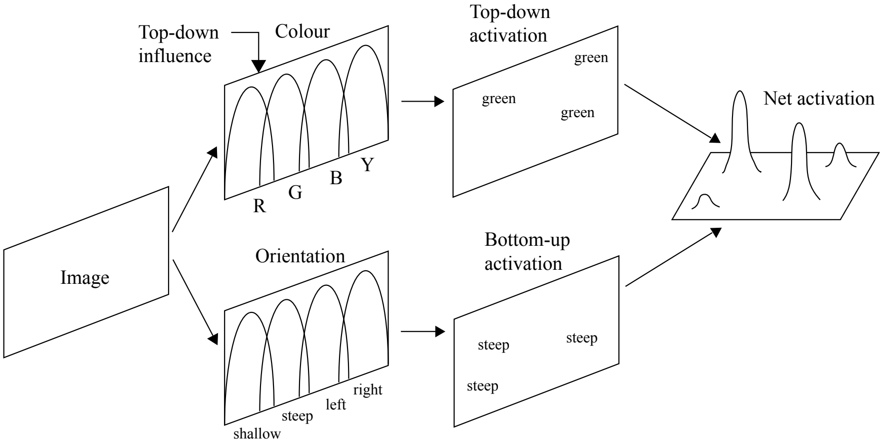

Figure 1.2. Schematic view of Wolfe's Guided Search Model (1989). An image is filtered into categories for different feature maps. Next top-down activation (determined by top-down influences such as task goals) and bottom-up activation (based on the salient part of the image) both mark regions of interest. A net activation map is build by combining these bottom-up and top-down activations. Attention is drawn to the highest 'hills' in this map.

as ends of one and the same continuum rather than a strict dichotomy (Rauschenberger & Yantis, 2006).

Examples of accounts that describe attentional selection on a continuum can be found in attentional engagement theory by Duncan and Humphreys (1989) and models such as Wolfe's guided search (Wolfe & Cave, 1989; Wolfe, 1994; Wolfe & Gancarz, 1996). The latter describes how an activation map is created based on the sensory image by two processes that work simultaneously. Both a top-down activation based on current task goals as well as a bottom-up stimulus-driven activation determine what part of the perceptual scene yields the strongest activation and thus will be selected for perception. By varying the strength of both of these components we can create a whole continuum of perceptual phenomena ranging from preattentive perception such as pop-out (Maljkovic & Nakayama, 1994) to instances where perception is clearly attentional and goal-driven (see Figure 1.2).

Although the models described above can perhaps give more insight on *how* parts of the sensory information are selected in perception, it still does not give insight in *how well* we can attentionally select this information.

Since attention is described in the models mentioned above, and in many other instances, as a selection process, we argue that attention can be seen as just another filter after early perceptual filters. This is close to what James already noted in 1890: "Everyone knows what attention is. It is the taking possession by the mind, in clear and vivid form, of one out of what seem several simultaneously possible objects or trains of thought." Thus, if we restrict ourselves to describing attention as a filter (mechanism), we can describe what the spatiotemporal properties of such a filter are and understand how well this attentional selection mechanism in perception works. However, since we describe a system that does not simply filter input from objects and events directly, but depends on both bottom-up and top-down processes, it has proven to be difficult to pinpoint what the spatiotemporal constraints of attentional selection mechanisms precisely are.

Hence, the psychophysical experiments described in this thesis will show new insights in the spatiotemporal limits of attentional selection under a diversity of conditions. A large part of the thesis will be concerned with thusfar less studied temporal aspects of these limits of attentional selection and how they affect perception. However, attentional selection occurs in both time and across space. Therefore, the last part of this thesis will be dedicated to new insights in spatial attentional selection as well.

The literature on attention (and its limits) is an extensive one. Though a lot of studies call the observations they find not necessarily (limits of) attentional selection, they do describe outcomes that are the result of basically similar attentional selection processes. This results in different limits in attentional selection being described as if they are limits of separate phenomena such as attentional blink (Shapiro, Raymond & Arnell, 1994), attentional capture (Jonides & Yantis, 1988; Hillstrom & Yantis, 1994), change blindness (Simons & Levin, 1997; Rensink, O'Regan & Clark, 2000), banner blindness (Benway & Lane, 1998).

We should, however, describe these phenomena in the light of their common denominators, which are the temporal and spatial limits of an attentional selection mechanism and study those limits. After we have given a short overview of what is known about temporal and spatial limits of attention, we will examine these limits in this thesis and thus investigate how well our top-down attentional control in perception is.

Temporal limits of attentional selection

When describing temporal limits of attention, we describe which and how many events and objects can be selected and separated by attention in time. In other words, multiple events and objects that enter an attentional selection mechanism, but are fused into a percept of a single event or object reflect that attention can no longer individuate events and objects temporally. Objects and events that fall within the same attentional time span will be selected as if stemming from one and the same source. This temporal attentional filter has its own properties and thus differs from early temporal filtering in our sensory system (Kelly, 1979). Such an attentional temporal filter can be seen as the resampling of sensory information that is already filtered by earlier systems. The question then arises, how long this attentional time span precisely is. Both for visual and audio-visual perception this time span is suggested to be quite long (Dixon & Spitz, 1980; Lewkowicz, 1996) and previous studies show that this time span might be as long as 500 ms. For example, in the attentional blink literature the ability to report a second target, presented after a first target in a stream of rapidly presented objects, is impaired up to 500 ms after the presentation of the first target (Raymond, Shapiro & Arnell, 1992; Duncan, Ward & Shapiro, 1994). However, this attentional blink starts at around 200 ms after the first target, which would suggest that within this 500 ms attention samples twice: once with a time span of 200 ms and once with a time span of 300 ms. This would suggest that objects and events that occur within the same time span of up to 300 ms are not separable by attention. Results from attentional capture research show a slightly different temporal attentional limit. Attention can be captured only within a time span of around 200 ms (Theeuwes, 2004; Theeuwes, De Vries & Godijn, 2003). Yet another line of research indicates that temporal filtering by attention might occur on a much finer scale, but depends on the number of features across which an attentional selection has to be made (Holcombe & Cavanagh, 2001). For instance, we can match different features alternating simultaneously in pairs, such as colour and orientation, up to alternation frequencies of 18.8 Hz. This means that observers can select simultaneously presented feature pairs from pairs that are separated by 27 ms, which is around an order of magnitude shorter than the 200-300 ms as indicated by attentional capture and attentional blink research. However, if the features are also separated in space, this temporal selection mechanism suddenly occurs at a much coarser scale. The time between alternations of feature pairs has to be around 160 ms, when colour and orientation alternate simultaneously, but on different spatial locations.

Temporal limits of attention have not only been studied within the visual modality. Recent research on audio-visual perception of synchronously occurring events has shown that, when these events are presented above alternation rates of 4 Hz, which comes down to 250 ms between events (Fujisaki & Nishida, 2005), they are no longer perceived as separate events. This limit was attributed by these authors to a central attentional mechanism, which they argue reflects the same limit as found in synchrony judgments in the visual domain (Johnston & Nishida, 2001; Moutoussis & Zeki, 1997; Nishida & Johnston, 2002).

Yet another area of research shows that the temporal limits of attention might be shorter than sampling every 100 ms. A study by Verstraten and colleagues (Verstraten, Cavanagh & Labianca, 2000) has shown that in an ambiguous apparent radial motion display observers are able to attentively track an object, if this object shifts to a new location every 83 ms. This means attention can disengage from its current location and select the same object at a new location within 83 ms. Note that other researchers (Horowitz, Holcombe, Wolfe, Arsenio & DiMase, 2004), using a similar paradigm, have found coarser limits (attention shifting circa every 200 ms), but this discrepancy is solved in Chapter 2 of this thesis.

Spatial limits of attentional selection

Results from different types of studies have yielded different outcomes with respect to the spatial resolution of attention (Intriligator & Cavanagh, 2001). Results from cueing paradigms, for example (Posner, 1978; Posner, Snyder, & Davidson, 1980; Sagi & Julesz, 1986) have shown that the attentional focus areas range from several arcminutes to around full hemifields. Studies on flanker interference (Eriksen & Hoffman, 1973; Eriksen & St. James, 1986) show variable results as well, which suggest an area of attentional selection that ranges from somewhere between 1 to 10 degrees visual angle.

In crowding studies (Bouma, 1970; 1973; Toet & Levi, 1992) the critical spacing, at which detection of the test item is at threshold, ranges from about one third of the eccentricity outside the fovea to a tenth of a degree inside the fovea. Intriligator and Cavanagh (2001) noted that the results of a number of counting studies show that items separated less than 5 arcminutes apart can not be counted at center of gaze. Because of to the large difference in the variable results of these different studies, Intriligator and Cavanagh (2001) set out to measure the spatial resolution of attention, defined as the resolution at which attention can still *individuate* objects. They found that the spatial resolution of attentional selection is inhomogeneous and also coarser than the visual spatial resolution. Spatial attentional resolution appeared to scale with eccentricity and was coarser in the upper compared to the lower visual field. More importantly, this resolution of attention does not depend upon the capacity of attention. The spatial attentional resolution will be the same, irrespective of whether attention has to select a single or multiple objects (see also Figure 1.3 for an example of such an attentional field).

Another important factor in spatial limits of attentional selection is the spatial distribution of attention across a visual scene. As the number of areas over which attention has to be divided increases, visual search performance, for example decreases (Gobell, Tseng, & Sperling, 2004).

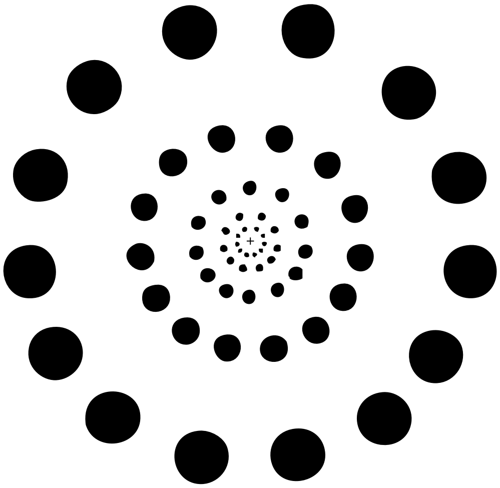

Figure 1.3. Viewing one's own attentional field (after Intriligator and Cavanagh, 2001). When viewing this stimulus such that it spans 30 degrees (which would be viewing it from approximately 14 cm) most observers will be able to view all items individually while fixating on the cross in the center. Viewing from a large distance or upside down will deteriorate the ability to individuate each object in this stimulus.

Interactions between perception and attentional selection

Not only does attention have spatial and temporal limits of its own, recent studies have shown that attention can actively modulate spatial and temporal resolution of earlier perceptual mechanisms as well. Carrasco and colleagues (2004) and Yeshurun and Carrasco (1999), for example, have shown that attention enhances spatial discriminability at any location of the visual field that is attended. Note that Tse et al. (2003) even find better change detection at locations opposite to a cued and thus attended area in the visual field. However, attention differentially altered temporal performance fields across the visual field (Carrasco, Giordano, & McElree, 2004). Information accrual without attention was fastest for targets on the horizontal meridian, intermediate for targets at the intercardinal locations, slow for targets on the vertical meridian, and slowest for targets at the upper location. These differences in temporal performance fields were eliminated when attending to these locations. In contrast, Yeshurun and colleagues actually find a decreased temporal performance due to attention being directed to certain locations in the visual field (Yeshurun & Levy, 2003; Yeshurun, 2004).

Outline of thesis

This thesis further examines the limits of attentional selection described above to investigate how well we can voluntarily control our perception, with a focus in chapters 2-4 on thusfar less examined temporal aspects of attentional selection. In Chapter 5 we report new insights in spatial aspects of attention as well. Thus, the experiments in all chapters together describe spatiotemporal attentional limits.

In Chapter 2 we describe how an apparent discrepancy between previous results concerned with the temporal limits of attentive tracking (Horowitz, Holcombe, Wolfe, Arsenio & DiMase, 2004; Verstraten, Cavanagh & Labianca, 2000) can be attributed to the timing of a certain signal, namely the *offset* signal of the tracked object. Attentive tracking is based on the phenomenon of attention-based motion perception, where selecting an object in an otherwise ambiguously radially apparent motion stimulus results in a percept of an unambiguously moving stimulus in clockwise or counterclockwise direction. This percept thus depends on the correct attentional selection of objects in time. The timing of the above mentioned offset signal turned out to be the critical determinant of the ability to attentively track an object in such an ambiguous motion display.

In Chapter 3 we focus on temporal attentional selection of rapidly alternating simultaneously presented features. In the visual domain the ability to integrate the correct features into alternating feature pairs temporally has been shown to critically depend on the number of features over which integration has to take place (Holcombe & Cavanagh, 2001). We hypothesized that sensory modality (i.e. the modalities through which information about an object or event reaches us) might just be another feature to an attentional selection mechanism. Therefore the temporal limit of attentional selection should decrease when selecting across modalities as compared to selecting within one modality. This decrease in the temporal limit of cross-modal integration compared to unimodal visual and unimodal auditory integration was indeed found in our experiments.

Instead of examining what attention can *select*, we can also examine how well attention can *ignore* irrelevant events and objects and thus select only the objects and events needed for performing a task. Thus, the same limits of attentional selection should apply to the ability to ignore certain irrelevant objects and events.

In Chapter 4 we examined recent accounts of attentional capture and temporal properties of events that lead to attentional capture (Lamy, 2005; Von Mühlenen, Rempel, & Enns, 2005). These accounts appear to contradict each other; both explain why certain temporal properties of irrelevant events determine whether our current attentional selection can be interrupted (in other words, capture our attention). We show an overall different pattern of results that was consistent over two experiments. Based on the results of Chapter 4, we could rule out one of the two accounts (Von Mühlenen et al., 2005) for certain. Our results indicate that it was not the expected occurrence of irrelevant events per se, as proposed by Lamy (2005), but the timing of the task-relevant target onset which determined whether observers could ignore an irrelevant event. Apparently, the timing of task-relevant events is yet another factor in determining the limits of attentional selection.

Finally, in Chapter 5 we describe a series of visual search experiments that show that observers can not ignore irrelevant non-targets in a large search display when the number of these non-targets is around 5% of the total number of non-targets. Though observers know and are instructed to ignore these irrelevant non-targets, search time is longest compared to conditions with either more or fewer irrelevant non-targets. This was found for multiple set sizes. Furthermore, the only measure correlated to search time in these experiments appeared to be the number of fixations. It might therefore be argued that these irrelevant non-targets decrease visual span maximally when 5% of a search display is irrelevant. Visual span can be defined as the area in which items are inspected during fixation. Apparently, attentional selection is limited by visual span. How precisely this visual span in turn is decreased by the relatively small number of irrelevant non-targets remains unclear.

This thesis can thus be divided into two sections; limits of attentional selection are explored by investigating attentional selection performance in Chapter 2 and 3, by determining what objects and events can be selected. In Chapter 4 and 5 the same attentional selection limits are examined, but more indirectly by showing what irrelevant events and objects are processed, while they should be ignored by attention. The focus thus changes from the selection of the task-relevant items in Chapter 2 and 3, to which irrelevant events and objects can not be ignored in Chapter 4 and 5.

In the final chapter I will shortly summarize these findings and discuss what the new insights in spatiotemporal limits in top-down attentional selection in perception are.

Chapter 2

Disengaging attention sets the temporal limit of attentive tracking

Abstract

At first sight, recent studies investigating the temporal limits of attentive tracking show contradictory outcomes. Attentively tracking an object in an ambiguous apparent motion display can have an upper limit of around 0.4 revolutions per second (rps; Horowitz et al., 2004) or 1 rps (Verstraten et al., 2000). Here, we demonstrate that this difference depends on presentation conditions: an important determinant for the temporal limit of attentive tracking appears to be the duty cycle. Tracking performance at high(er) rates decreases to chance with increasing duty cycle, while at low rates duty cycle hardly has an effect on performance. Results are discussed in terms of (dis)engagement of attention.

Introduction

In 1912 Wertheimer showed that alternating two frames, each containing a cross in which the cross is rotated 45 degrees in one frame relative to the same cross in the other frame, results in perceiving rocking motion when viewed passively. The perceived direction is ambiguous, since clockwise and counterclockwise motion is equally probable. Attentively tracking one of the spokes of the cross, however, results in a clear apparent motion percept which is called attention based apparent motion (Verstraten, Cavanagh, & Labianca, 2000). This motion percept can also be achieved for ambiguous continuous stimuli like radial gratings (Cavanagh, 1992), a phenomenon known as *attention based motion perception*.

Recently, several researchers have addressed the question about the temporal limits of attentive tracking. Horowitz and colleagues (Horowitz, Holcombe, Wolfe, Arsenio, & DiMase, 2004) found an upper limit of attentive tracking around 2.5 Hz, whereas Verstraten et al.'s data (2000) suggested a 5-7 Hz limit (Horowitz et al. express their limit in terms of duration of one frame interval. Their limit lies around a minimum duration of 200 ms, while Verstraten et al. (2000) find a limit of around 70 to 100 ms). Horowitz et al. (2004) explain this difference by suggesting that the temporal limit of attentive tracking reported by Verstraten and his colleagues (2000) is not a limit of attentive tracking, but rather is a limit of object continuity. In their discussion they suggest that in Verstraten's paradigm observers are merely indexing objects (using FINST (Pylyshyn, 1989) or "object files" (Kahneman & Treisman, 1984)) rather than tracking them using attention. They suggest that, though information about the features is not available, indexing is enough to keep track of an object and is faster than attentional pursuit; only the spatiotemporal history of an object's index needs to be kept track of. However, the spatiotemporal history of the index of an attentively tracked object is ambiguous, whereas the stimuli in the studies on object continuity (Pylyshyn, 1989; Kahneman & Treisman, 1984; Kahneman, Treisman, & Gibbs, 1992) are stationary or move unambiguously. An object and thus its index in both attentive tracking paradigms could have moved either clockwise or counterclockwise between frames with equal probability. Constant attentional selection is thus needed for tracking an object in both Verstraten et al.'s (2000) and Horowitz et al.'s (2004) experiments.

If it is not attentional selection itself, what then constitutes the difference between the temporal limits found by these researchers? Here, we suggest that the differences can be explained by the ability of attention to disengage from one location of a tracked object and engage to the next location of that same object. It is known from both eye movement literature (Fischer & Ramsperger, 1984) and from research on express attentional shifts (Mackeben & Nakayama, 1993) that shifting from one location to another can be faster when previous fixated or attended objects disappear before they reappear elsewhere. In Horowitz et al.'s (2004) experiment offset of one frame and onset of the next coincided, i.e. the duty cycle of each element was 100%. Duty cycle is defined as the duration of one frame expressed as a percentage of the total duration of one frame and a following blank interval. In Horowitz et al.'s (2004) experiment no blank interval was present. The onset of the next frame of placeholders in their attentional pursuit condition triggers attention to make a shift to that next location. Though Horowitz and colleagues (2004) state that attention does not leave the tracked object, attention still needs to shift from location to location. In contrast, Verstraten et al. (2000) did use blank intervals between two consecutive steps of a tracked object. The duty cycle of a tracked object was 40% in their experiment. Thus, the offset of the tracked object triggered attention to disengage and shift to a next location *before* the onset of the tracked object at the next location. Here, we therefore manipulate duty cycle systematically.

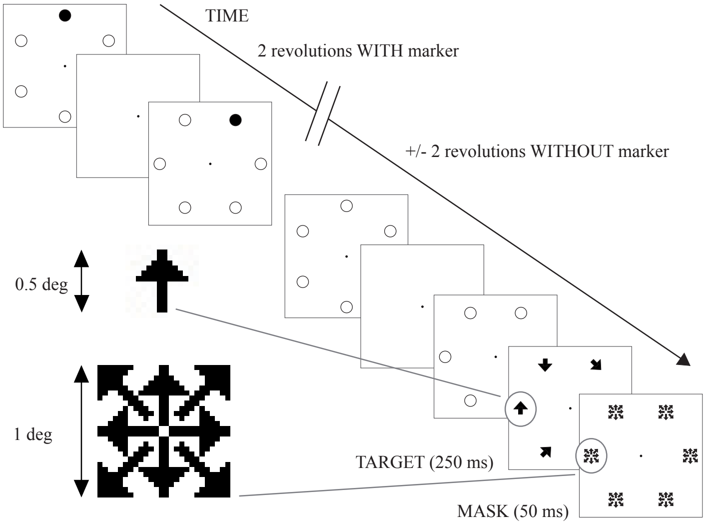

Figure 2.1. Schematic overview of the stimulus.

Methods

Participants

Four observers (3 naïve as to the purpose of the experiment) voluntarily participated in the experiment. All had normal or corrected to normal vision. One observer was the first author (JB).

Stimuli

Stimuli were created and presented using Matlab® 5.2.1 and the Psychophysics Toolbox extensions (Brainard, 1997; Pelli, 1997) on a Macintosh G4 computer. Stimuli were presented on a 19 inch Iiyama Pro Vision Master 454 monitor set to a resolution of 640 by 480 pixels and a vertical refresh rate of 200 Hz yielding 5 ms timing precision. Participants were placed in a head and chin rest such that the distance to the screen was 86 cm.

Two circular arrays of six evenly spaced discs were alternated in space and time, separated by a blank interval (except for one condition, see below). One of the two frames containing discs was rotated such that a disc in one frame was located exactly between two other discs in the other frame. The center of the resulting 12 locations a disc in the stimulus could occupy were therefore 30 degrees separated. Discs were white (luminance: 133 cd/m^2) on a grey (luminance: 15.1 cd/m^2) background and had a diameter of 0.5 degrees. The radius of the circular array on which the discs were placed was 4 degrees and was centered around a 4.2 arc minutes black (luminance: 0.26 cd/m^2) fixation point.

Manipulating the duration of the blank interval resulted in different duty cycles of the discs: 25, 50, 75 and 100%. In the latter case there was no blank interval between the presentations of the two circular arrays, similar to Horowitz et al's (2004) attentional pursuit of placeholders. The rate at which these circular arrays alternated was manipulated as well in such a way that discs could travel either at 0.52 or 1.04 revolutions per second (rps). In terms of interval duration (duration of one frame plus following blank interval) this is 160 or 80 ms, respectively. The lower of these rps values approximates the temporal limit found by Horowitz et al. (2004) the other rate is close to the limit found by Verstraten et al. (2000).

Task and procedure

Tracking in the resulting eight conditions (two levels of rps x four levels of duty cycle) was tested using a response method similar to Horowitz et al. (2004). One of the white discs (random per trial) in the array started out black. This black disc shifted in one of the two possible directions (clockwise or counter-clockwise, random per trial). After two full revolutions the black disc turned white and the observer's task was to attentively track this disc. After approximately two more revolutions the alternation of the disc arrays was replaced by probe arrows, presented (0.5 degrees width for 250 ms) at the six locations of the discs. These probes were subsequently masked by six 1 degree stimuli consisting of a constellation of all arrow orientations. These masks were presented for 50 ms (see Figure 2.1 for a schematic overview of the sequence). After the presentation of the masks subjects had to report the direction of the arrow presented at the location of the tracked disc using the number pad of the keyboard. Arrows could point in one of eight directions. The arrows at the locations of the non-tracked discs always pointed in different directions than the target arrow. Each combination of rps and duty cycle was first practiced in a block of 80 trials by each observer. Thereafter, each observer was tested in two blocks of 25 repetitions of each condition (200 trials per block).

Results

The results in Figure 2.2 (upper left panel) show the mean of the 4 average proportions correct, calculated per subject (other panels). Individual averages are based on the 50 repetitions per point from the two experimental blocks. For each of the 4 subjects these results show that at the lower revolution rate of the discs duty cycle has little effect on

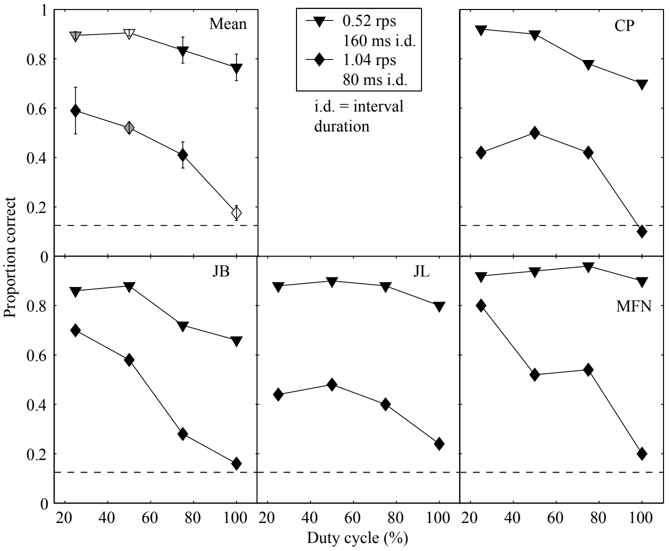

Figure 2.2. Upper left panel: average proportion correct (4 subjects, errorbar = +/- 1 SEM). Gray markers indicate conditions with same duration of last frame before the probes (40 ms). Analoguously white markers indicate conditions where this duration is 80 ms. Other four panels: individual proportions correct (50 repetitions per point). The dashed horizontal lines indicate chance performance.

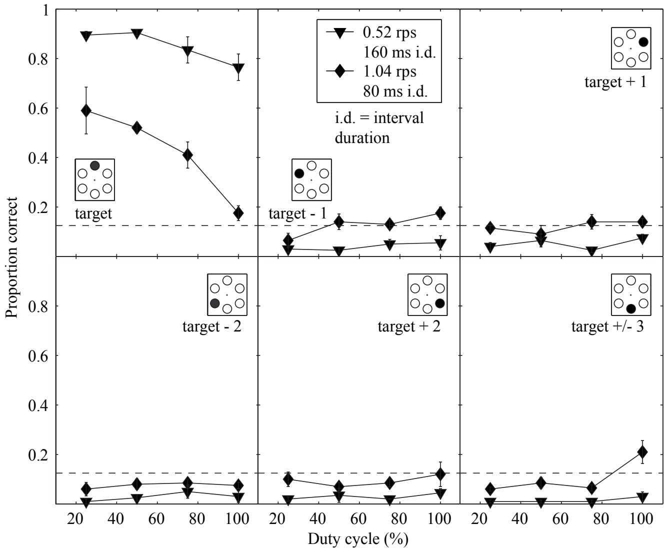

Figure 2.3: average proportion correct (4 subjects, errorbar = +/- 1 SEM) for all discs (target location (upper left panel) and surrounding non-target locations). The dashed horizontal line indicates chance performance.

tracking performance. Subjects can report the orientation of the arrow target well above chance performance (0.125). The mean proportion correct responses in the 0.52 rps (160 ms interval duration) condition is 0.90 at 25% and 0.77 at 100% duty cycle. At the higher of the two revolution rates (1.04 rps, 80 ms interval duration) the proportion correct depends much more on duty cycle and drops from 0.60 at 25% duty cycle to 0.17 at 100% duty cycle, which is around chance performance. The performance of 60% is expectedly somewhat lower than performance of 66.7% in Horowitz et al. (2004) due to smaller interval duration in the current experiment. The smallest interval duration which yielded 66.7% performance in their experiments was 107 ms.

With increasing duty cycle the last frame containing discs, before the frame containing the probes, is presented longer, which possibly results in stronger forward masking of the probes by the discs. However, temporal duration of this last frame is the same at the two revolution rates at some points (e.g. 50% duty cycle at 0.52 rps and 100% duty cycle at 1.04 rps both result in 80 ms duration of the last disc frame), while performance clearly differs for the two revolution rates at these points (marked grey and white in Figure 2.2). Thus, it is unlikely that the effect of increasing duty cycle on performance can be attributed to forward masking.

It could be argued that when tracking becomes more difficult at higher revolution rates of the stimulus, subjects ‘miss’ or ‘skip’ an attentional step and are therefore not responding to the disc to be tracked but to the disc at a location before or after the target location. A second analysis shows that subjects did not systematically report the arrows’ orientation of one of the other five disc locations. The mean proportion correct stays around chance level (1/8, eight directions to possibly respond to) for all other locations. This can be seen in Figure 2.3, where for the target location and all other five locations (expressed as target location plus or minus 1, 2 or 3), the response for all duty cycle and rps combinations stays around the dashed line indicating chance performance. Tracking is just impossible with larger duty cycles at the high revolution rate.

Discussion

In the current experiment we have shown that the temporal limit of attentive tracking (or attentional pursuit) of an object in an ambiguous apparent motion display is determined not by the revolution rates of objects in that display per se. Rather, it is the duration of a blank interval between the two frames that make up the display that determines whether objects can be tracked at a specific revolution rate or not. More specifically, subjects in our study can track an object at 0.52 revolutions per second (rps), or 160 ms interval duration, independent of the duty cycle (the ratio between duration of one frame with elements and the duration of the following blank interval). For a higher rate (1.04 rps, 80 ms interval duration), subjects can only track an object when the duty cycle of the object is 50% or smaller. At 100% duty cycle tracking performance drops to 17% which is near chance performance (12.5%). Recently, it was reported that attentional pursuit has an upper temporal limit of around 0.42 rps, or 200 ms interval duration, (Horowitz et al., 2004) as opposed to previous results that showed attentive tracking was possible up to revolution rates of around 1 rps, around

80 ms interval duration (Verstraten et al., 2000). Horowitz et al. (2004) explained this difference by suggesting Verstraten et al. (2000) had measured object continuity rather than attentional pursuit. They reasoned that in the latter of the two experiments subjects could just 'index' the to-be-tracked object (Kahneman & Treisman, 1984, Pylyshyn, 1989) and update the spatiotemporal history of that index by means of low-level motion processing. In contrast to these studies (Kahneman & Treisman, 1984, Pylyshyn, 1989), where stimuli were stationary or moved unambiguously, motion of identical objects in both Horowitz et al.'s (2004) and Verstraten et al.'s (2000) paradigm is ambiguous. As argued in the introduction, the spatiotemporal history of the indices belonging to those objects is ambiguous as well. Thus, in both studies constant attentional selection of an object is needed to resolve the ambiguous motion of that object.

With the current results we can now explain the difference found in these studies simply in terms of the temporal layout of the stimulus configuration. In their attentional pursuit condition Horowitz and colleagues (2004) alternated the two frames with six elements without blank intervals, while in the study by Verstraten et al. (2000) a blank interval was interleaved with the frames which contained the elements. Thus, there was an interval between the offset of an object in one location and the onset of that object in the next location in the latter study. Studies on eye movements and attentional shifts show that moving the eyes or attention to a next location is speeded when the previous fixated or attended object is removed (Fischer & Ramsperger, 1984, Mackeben & Nakayama, 1993). This is explained in terms of attention being able to disengage before onset of the next object and thus speeding up (attentional) pursuit or detection of objects. Tam and Stelmach (1993), on the other hand, attribute shorter saccade latencies due to a (temporal) gap between offset of a dot and onset of a target dot not solely to disengagement of attention, but suggest either solely ocular or ocular-attentional disengagement. Klein, Taylor and Kingstone (1995) have reinterpreted these gap effect results as being solely ocular based. Alternatively, Danckert and Maruff (1997) reinvestigated this gap effect using the original paradigm of covert orienting of visual attention (COVAT) by Posner (1980). Since a gap effect was not found, they concluded that covert attentional processes can only be facilitated when ocular systems are in a disengaged state. In the current experiment, in Danckert and Maruff's (1997) terminology, an ocular system is continuously engaged on fixation, while attention is shifting through multiple locations. If attentional systems are overruled by ocular engagement, we should not have found an effect of our duty cycle manipulation. We therefore suggest that covert attentional mechanisms operate independently from ocular systems, making attentional (dis)engagement possible while ocular systems are engaged.

A further question is whether the disappearance of objects (the offset) acts as a cue for attention to shift to another location or that disappearance works differently in determining the speed of attention shifts. Mackeben and Nakayama (1993) already tested this 'readiness/alerting hypothesis' and found that removing an object before target onset shows speeding of attentional shifts, whereas using other cues like a change in fixation mark or shortly increasing brightness of the screen does not. Apparently, element offset does not act in the same way as a cue. It is the timing of disappearance of an object that determines the rate of attentional shifts.

In sum, this study shows that the temporal limits of attentive tracking depend on duty cycle, which facilitates attentional disengagement independently from ocular systems.

Chapter 3

Matching auditory and visual signals: is sensory modality just another feature?

Abstract

In order to perceive the world coherently, we need to integrate features of objects and events that are presented to our senses. Here we investigated the temporal limit of integration in unimodal visual and auditory as well as cross-modal auditory-visual conditions. Participants were presented with alternating visual as well as auditory stimuli and were asked to match them either within or between modalities.

At alternation rates of about 4 Hz and higher, participants were no longer able to match visual and auditory stimuli across modalities correctly, while matching within either modality showed higher temporal limits. Manipulating different temporal stimulus characteristics (stimulus offsets and/or auditory-visual SOA) did not change performance.

Interestingly, the difference in temporal limits between cross-modal and unimodal conditions appears strikingly similar to temporal limit differences between unimodal conditions when additional features have to be integrated. We suggest that adding a modality across which sensory input is integrated has the same effect as adding an extra feature to be integrated within a single modality.

Introduction

A coherent representation of objects and events in the world around us depends on integration of multiple features thereof across space and in time. Our brain appears to accomplish this task almost effortlessly, although integration of features, especially when they originate from different dimensions, has spatial as well as temporal limits.

In visual perception, temporal integration limits are usually explored by presenting rapidly alternating feature pairs. Typically, subjects are asked to match features that occur simultaneously. A study by Holcombe and Cavanagh (2001), for example, shows that correctly matching simultaneous alternations in stimulus' orientation and colour is limited to alternations up to 18 Hz. Moreover, this limit decreases to about 3 Hz when these features are spatially separated. Apparently, as features are separated over more than one dimension (space and time vs. time alone), the temporal limit of integration changes dramatically.

In audition, models of temporal integration have been divided into two classes favouring different integration times. One line of research suggests that all auditory input that falls within a window of about 200 ms is integrated (5 Hz), while results based on gap detection and discrimination favour integration times of about 5 ms (200 Hz), the so-called "resolution-integration paradox" (De Boer, 1985; Green, 1985). Viemeister and Wakefield (1991) showed that increasing the noise level of a 100 ms interval between two sound pulses did not decrease detectability of the individual pulses, which argues against a large temporal window of integration. Moreover, even when the two pulses were separated by only 5 ms, they still appeared to be processed independently.

Clearly, auditory and visual perception have different temporal limits of integration. Yet, we often need to integrate across senses as well. Here we aim to determine the temporal limit of integration when we integrate across these two modalities. More specifically, we address the question whether the auditory-visual temporal integration limit differs from the limits set by either the visual or auditory modality.

So far, auditory-visual temporal integration research has typically focused on phenomena like auditory capture (Fendrich & Corballis, 2001; Gebhard & Mowbray, 1959; Recanzone, 2003; Shipley, 1964), synchrony judgments (Fujisaki, Koene, Arnold, Johnston & Nishida, 2006; Fujisaki & Nishida, 2005; Fujisaki, Shimojo, Kashino & Nishida, 2004; Noesselt, Fendrich, Bonath, Tyll & Heinze, 2005) and temporal order judgments (Keetels & Vroomen, 2005; Morein-Zamir, Soto-Faraco & Kingstone, 2003; Vroomen, Keetels, de Gelder & Bertelson, 2004). In these paradigms the auditory and visual signals are not presented synchronously. Subject's sensitivity to this auditory-visual asynchrony is measured, which is a measure of segregation. However, even though integration of specific features without segregation of others cannot exist, the temporal limit of *segregation* of auditory-visual signals is not necessarily the same as the temporal limit of *integration*. The same separate sensory events might be perceived asynchronously in one task, while they might be integrated into a single percept in another. Moreover, depending on whether the task is to integrate or segregate sensory input, spatial location, for example, can have opposite effects on temporal limits. As mentioned, matching in Holcombe and Cavanagh's (2001) unimodal task has a higher temporal limit when features alternate at the same spatial location, while in auditory-visual temporal order judgments the just noticeable difference (JND) between auditory-visual signals is smaller (i.e. there is a higher temporal limit) when features are spatially separated (Keetels & Vroomen, 2005). With this possible task-dependent effect on perceiving sensory events synchronously in mind, we specifically investigate the limits of temporal integration by using a matching paradigm in the visual and auditory domain as well as across these modalities.

Experiment 1 Cross-modal versus unimodal matching

This experiment aimed at finding the temporal limit of matching auditory and visual input. The two sensory systems we tap into here both have separate limits of temporal integration (Holcombe & Cavanagh, 2001; Viemeister & Wakefield, 1991). The question arises whether one of these systems forms a bottleneck in auditory-visual integration, or whether the cross-modal integration limit differs from either one. If the latter is the case it makes sense to expect that limit to lie below the integration limit of either modality.

Methods

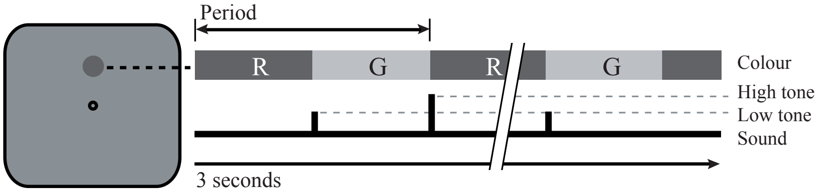

Figure 3.1a. Schematic overview of a trial. Example of red (R) / high – green (G) / low presentation. The grayscale bar (dark gray = red, light gray = green) and pulse time line together indicate that a high tone is presented at onset of each red disc.

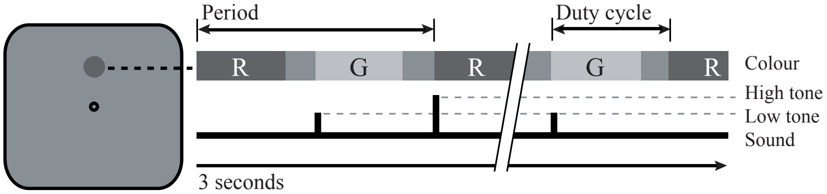

Figure 3.1b. Schematic overview of a trial. Example of red (R) / high – green (G) / low presentation in a 75 % duty cycle trial. The grayscale bar (dark gray = red, light gray = green, intermediate gray = coloured disc off) and pulse time line together indicate that a high tone is presented at onset of each red disc.

Participants

Four subjects (two naïve as to the purpose of the experiment) participated in the experiment. All participants had normal or corrected to normal vision and no known hearing problems.

Stimuli and procedure

Stimuli were created and presented using Matlab® 5.2.1 and the Psychophysics Toolbox extensions (Brainard, 1997; Pelli, 1997) on an Apple G4 computer. The visual stimuli were presented on a 19 inch Iiyama Pro Vision Master 454 monitor, set to a resolution of 640 by 480 pixels and a refresh rate of 200 Hz, yielding 5 ms timing precision. The auditory stimuli were presented through Sennheiser HD-250 linear II headphones, connected directly to the audio-port of the computer. The participants were seated 86 cm from the monitor and their head was supported by a chin and forehead rest.

Auditory-visual conditions

The visual stimulus consisted of isoluminant (35.6 cd/m^2) red and green discs spanning 0.75 degrees of visual angle. The discs alternated in colour with different frequencies. The rate of alternation was varied per trial. Alternation rate was defined as the inverse of the period from one onset of a coloured disc (e.g. red) to the onset of the next disc with the same colour (see Figure 3.1a). Five alternation rates were used: 1, 1.9, 3.1, 4.2 and 5 Hz. For each trial, the coloured discs alternated for a total of 3 seconds. The discs were presented 3 degrees above or below a 0.5 degrees fixation spot (black/grey bull's eye) on a grey background (15.1 cd/m^2). The location of the visual stimulus was varied to control for different temporal performance fields that have been demonstrated in visual perception (Carrasco, Giordano & McElree, 2004; Yeshurun & Levy, 2003), and which might differentially affect auditory-visual integration as well. Results from the studies mentioned above show a temporal performance advantage in the lower visual field.

Two sounds (a high-pitched tone of 2 kHz, and a low-pitched tone of 1 kHz, 78.9 dB and 77.9 dB sound pressure level (SPL), respectively) were presented simultaneous with the onset of the red and the green disc. Each tone had a duration of 25 ms and contained no spatial information. Tone duration was kept constant, since changing duration of tones along with alternation rate yielded perceptually different sounds. When a high-pitched tone was presented at onset of the red disc, a low-pitched tone was presented at the onset of the green disc. Physical onset differences between sound and colour stimuli ranged from 1 to 1.88 ms, which coincided with the vertical build-up of the screen. Which tone pitch was presented simultaneous with which colour was randomly varied between trials.

Auditory-visual procedure

The subject's task was to fixate the bull's eye, attend to both sound and colour of the stimulus, and after the stimulus had ended to indicate which sound accompanied the red disc. When a high-pitched tone was perceived simultaneous with the red disc, participants pressed the up arrow key on the computer keyboard. When a low-pitched tone was perceived, they pressed the down arrow key. To prevent participants from matching the correct sound to colour solely based on their first or last simultaneous onsets, the first and last onsets of the coloured discs were not accompanied by sound. Per trial the number of first colour onsets without sound was randomly varied (either one or two colour onsets) as was the number of last colour onsets without sound (again either one or two colour onsets). Each participant was presented with a block of 150 trials (5 alternations rates, 30 times repeated in random order).

Unimodal visual conditions

In the unimodal visual conditions, the red/green alternating disc was identical to that in the auditory-visual conditions, but the auditory signal was replaced by a visual stimulus. A second set of smaller (0.25 deg visual angle) alternating black and white discs was superimposed on the coloured discs. As in Holcombe and Cavanagh's visual matching paradigm (2001), a second feature (contrast polarity of the smaller disc) was superimposed on the coloured discs. The participant's task was to indicate whether the white or the dark disc coincided with the red disc. Otherwise, the procedure was identical to the auditory-visual procedure.

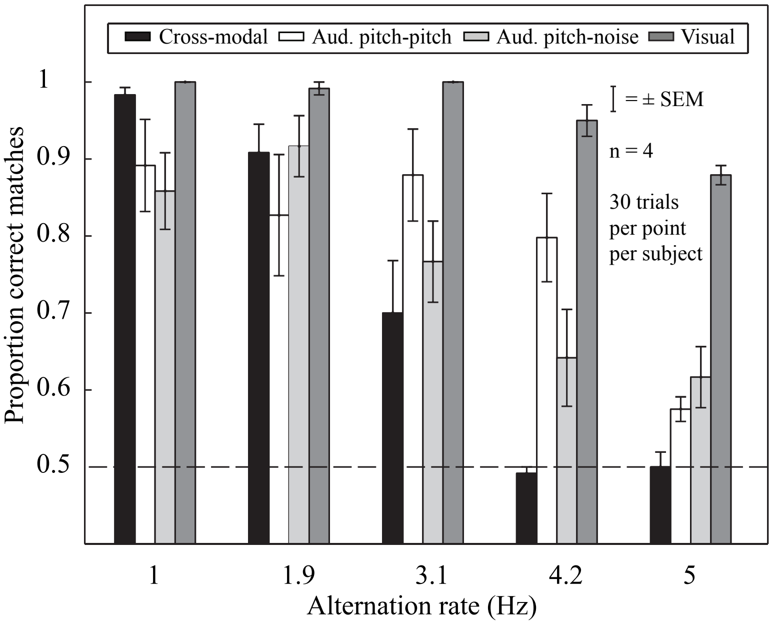

Figure 3.2a. Average proportion correct matches for the different alternation rates in cross-modal (black bars), visual (dark gray bars) and auditory (white and light gray bars) matching conditions. Error bars are +/- 1 S.E.M. The dashed line indicates chance performance.

Unimodal auditory conditions

For the unimodal auditory conditions, the two sounds (a high-pitched tone of 2 kHz, and a low-pitched tone of 1 kHz) were presented monaurally for 25 ms. A second set of alternating tones with different pitch (3.3 and 4.3 kHz, 73.4 dB and 70.2 dB SPL, respectively) was presented to the other ear. The participant's task was to indicate whether the two higher pitched tones were presented simultaneously (i.e. 2 kHz in one ear and 4.3 kHz in the other), or whether the high-pitched tone of one pair coincided with the low-pitched tone of the other. Otherwise, the procedure was identical to the auditory-visual procedure. It could be argued that by presenting to be matched auditory stimuli diotically, we introduce spatial separation of signals, thus making it harder to match simultaneous presented sounds. A second auditory matching paradigm was designed to test this. The high and low (2 and 1 kHz) tones were presented in both ears and were presented simultaneous with two alternating types of band-passed noise ('high' band-pass: 4.4-4.5 kHz, 'low' band-pass: 0.4-0.5 kHz). Pitch-noise ratio was set to 9:1. This prevented the tones from being drowned out by the band-passed noise.

Again, the participant's task was to indicate whether the high-pitched tone was presented simultaneously with the 'high' noise, or whether the high-pitched tone coincided with the 'low' noise. The sound pressure level (SPL) of the pitch-pitch combinations was 79.35 dB (SD = 0.64 dB) and 76.02 dB (SD = 0.69 dB) in the pitch-noise paradigm. Thus, participants could not determine a match based on SPL differences.

Results and discussion

Figure 3.2a shows the results collapsed across visuo-spatial conditions for both the unimodal and cross-modal conditions, since the effect of spatial location (upper versus lower visual field) as reported by Carrasco et al. (2004) and Yeshurun and Levy (2003) was not apparent from our data. Both in cross-modal and unimodal visual conditions the difference between presenting the visual stimuli above or below fixation was not significant ($F_{\text{cross}(1,\,8)} = 0.004$, $p = 0.951$ and $F_{\text{vis}\,(1,\,8)} = 5.16$, $p = 0.053$, see Figure 3.2b). It is immediately clear from Figure 3.2a that the temporal limit for the unimodal visual matching lies well above 5 Hz (the highest alternation frequency used) as, on average, performance is still above 85% correct at this frequency. The unimodal auditory matching performance is generally worse than that for the visual conditions, demonstrating that this is a more difficult task. This is reflected in the larger variance in performance at the lower frequencies; two participants started out with a performance of 80% correct, while the other two showed 98% correct at 1 Hz for pitch-pitch matches. For pitch-noise matches these percentages were 78% and 93%. This auditory matching performance declines to 58% correct for pitch-pitch matches and 61% for pitch-noise matches at 5 Hz.

Mann-Whitney U-tests (proportion correct is assumed not to be normally distributed) of the 5 Hz data show that unimodal matching performance still differs significantly ($U = 0$, $p < 0.05$ for visual, and $U = 0$, $p < 0.05$ for both auditory pitch-pitch and pitch-noise conditions) from chance performance, while cross-modal matching does not ($U = 8$, p

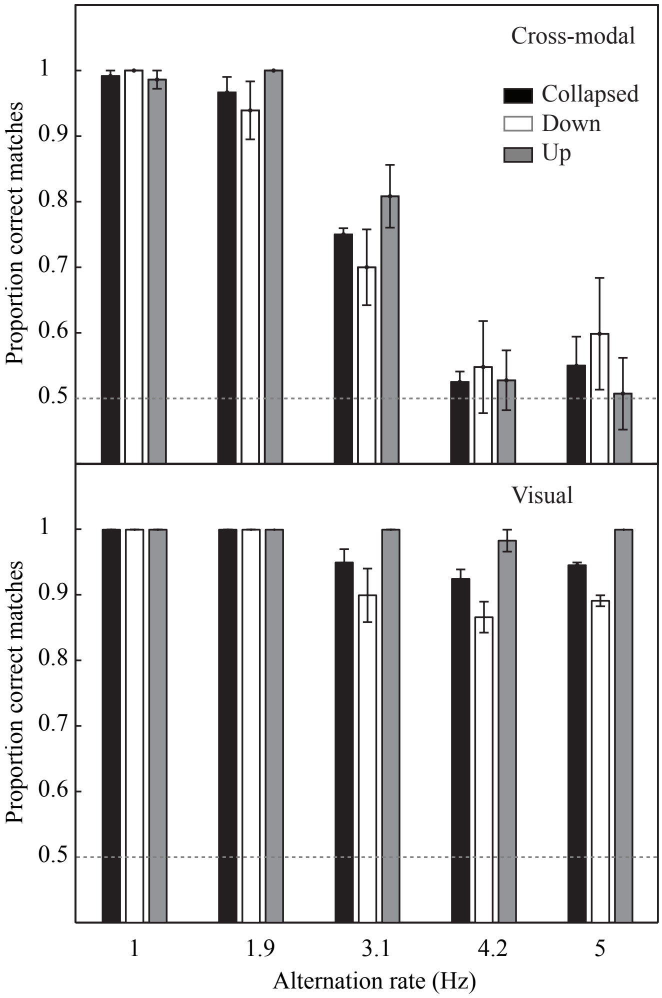

Figure 3.2b. Average proportion correct contrast polarity-colour (visual, lower panel) and sound-colour (cross-modal, upper panel) matches from experiment 1 for the different alternation rates: collapsed for spatial location (black bars) and separated for below (white bars) and above (gray bars) fixation. The dashed line indicates chance performance.

=1): as alternation rate increases, performance decreases from above 95% correct at 1 Hz alternation to chance level at alternation rates of 4.2 Hz and higher (dark bars). At these frequencies observers were no longer able to match sound and colour correctly. This is consistent with a pilot study using more alternation rates (1 to 5.5 Hz with steps of approximately 0.5 Hz), which resulted in the same limit (see Figure 3.2c).

From these results it is clear that neither temporal integration limit of the visual modality, nor that of the auditory modality forms a bottleneck for the auditory-visual matching performance. Note that the same visual and auditory stimuli were used, where possible, in cross-modal and unimodal conditions. Cross-modal matching shows a limit around or below 4 Hz, while the limit of unimodal matching lies above 5Hz. The fact that the unimodal matching limit lies above 5 Hz is in line with the results mentioned above (Holcombe & Cavanagh, 2001; Viemeister & Wakefield, 1991).

Experiment 2: Transient on- and offsets

In the first experiment, the auditory stimulus was a short transient stimulus with clear on- as well as offsets. The coloured discs, however, were isoluminant (even though they were not isoluminant to the background). The transition of one colour to the other (on- and offset at the same time) coincided with the onset of the auditory stimulus. A possible explanation for the low performance in the cross-modal conditions is that the onset of the auditory stimulus sometimes becomes associated with the offset instead of onset of one particular colour. Therefore, colour (and luminance) offsets were introduced into the visual stimulus in the second experiment, which only tested cross-modal conditions.

Methods

Participants and procedure were identical to those in the auditory-visual conditions of experiment 1. The auditory stimuli were only 5 ms in duration. The visual stimulus differed in the following way: interleaved with the red-green alternating disc, a blank interval was presented to investigate the effect of colour offset. The duration of this blank interval was varied as a function of the duty cycle of a coloured disc in five possible conditions: 100 % duty cycle (no blank interval, red changed directly to green; similar to the auditory-visual condition of experiment 1), 75 %, 50 %, 25 % duty cycle (see Figure 3.1b for an example of a 75 % duty cycle trial) and a fixed 10 ms disc presentation (each coloured disc was presented for a fixed duration of 10 ms). Each participant was presented with three blocks of 250 trials (5 alternations rates times 5 offset timing conditions, 10 times repeated per block in random order).

Results and discussion

Figure 3.3a shows that as alternation rate increases, performance of auditory-visual matching decreases, as was already shown in experiment 1. The shorter duration of the auditory stimulus did not influence the results, as is apparent from the dark bars (100 % duty cycle condition, which is a replication of the cross-modal results from experiment 1). For all duty cycle conditions the proportion correct matches decreases from about 90% at the lowest alternation rate to chance level at alternation rates of 4.2 Hz and higher. This is independent of the used duty cycle of the visual stimulus. The one exception is the condition where the visual stimulus had a fixed presentation duration of 10 ms (most similar in duration to the 5 ms auditory stimulus across all alternation rates used). For this condition performance was even worse. At 1.9 Hz alternation the performance is already at chance level. An additional experiment at a fixed alternation

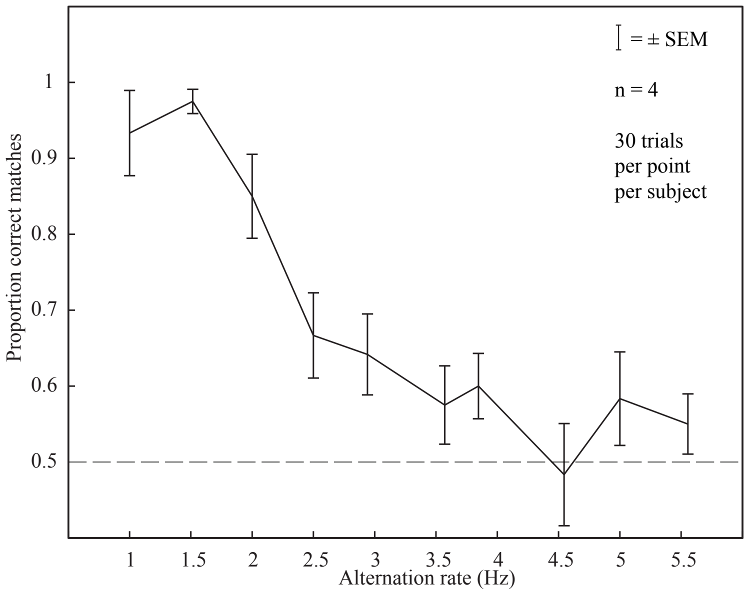

Figure 3.2c. Pilot data: Average proportion correct sound-colour matches. Here, more alternation rate levels were tested. Stimuli and procedure were the same as the 50% duty cycle condition of experiment 2. The dashed line indicates chance performance.

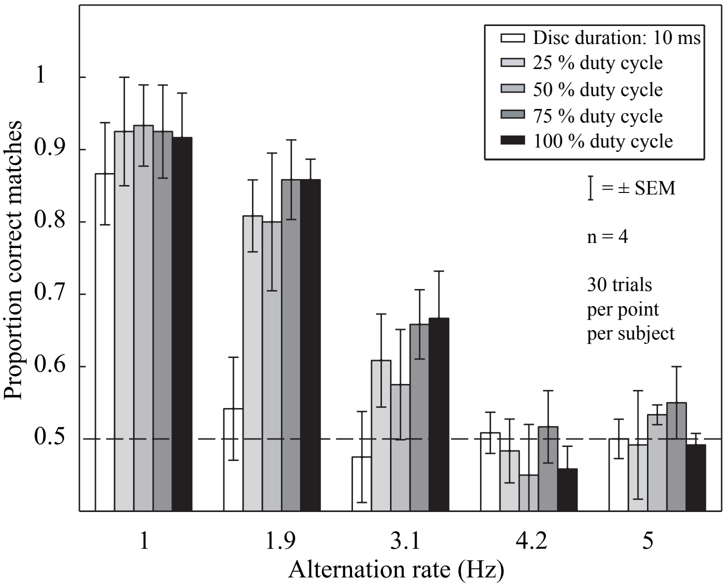

Figure 3.3a. Average proportion correct sound-colour matches for the different alternation rates. Error bars are +/- 1 S.E.M. in both panels. The dashed line indicates chance performance.

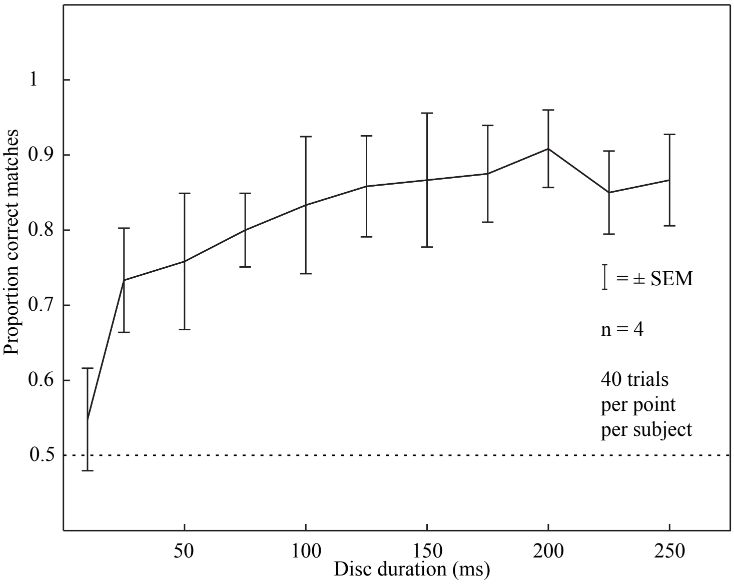

Figure 3.3b. Average proportion correct sound-colour matches at an alternation rate of 2 Hz, for several presentation durations of the visual stimuli. The task for the observers was the same as in experiment 2. The dashed line indicates chance performance.

rate of 2 Hz, with varying presentation duration of the visual stimulus (10 ms, and 25-250 ms with 25 ms steps) but otherwise identical to the experiment above showed that only at the lowest (10 ms) presentation duration this strong decrease in performance is apparent (see Figure 3.3b). At 25 ms presentation duration, performance is on average still above 70% correct. It seems plausible that at 10 ms presentation duration the discriminability (or visibility) of the visual stimulus is somewhat impaired, making the cross-modal matching task at 1.9 Hz already more difficult.

Nonetheless, the overall results of this experiment show that adding a transient visual offset signal (and thus additional temporal markers to which the auditory signal can be matched or mismatched) does neither facilitate nor impede auditory-visual matching.

Experiment 3: Perceptual delay

Several studies have demonstrated a perceptual delay between auditory and visual signals (Fujisaki et al., 2004; Keetels & Vroomen, 2005; Lewald & Guski, 2003; Vroomen et al., 2004; Zampini, Shore & Spence, 2005). The size of this perceptual delay ranges from between 7-10 ms (Fujisaki et al., 2004; Keetels & Vroomen, 2005) to 50–100 ms (Lewald & Guski, 2003), which is an order of magnitude greater. Any perceptual delay between sound and colour might account for decreasing matching performance when alternation rate increases in the previous experiments. Therefore, in this final experiment we investigated whether low performance in the previous experiments is confounded by perceived asynchrony between visual and auditory stimuli. To this end physical sound and colour onset synchrony was systematically varied.

Methods

Stimuli and task were the same as in the 100 % duty cycle condition of the second experiment. Only two alternation rates were used: 1.25 and 2.5 Hz. The temporal differences in sound and colour onset were systematically varied over a range of –189 to 189 degrees phase shifts in 29 equal steps. Negative onset asynchronies meant sound onsets preceded colour onsets (see Figure 3.4). Three subjects (two naïve) were tested in six blocks of 290 trials (29 phase shifts times 2 alternation rates, repeated 5 times per block).

Results and discussion

The proportion of the high-pitched tone being perceived with red in Figure 3.5a is plotted as a function of the presented sound-colour onset asynchrony. Since the subjects are asked to report which (high- or low-pitched) tone occurs simultaneously with the colour red, a proportion of 1 indicates that a high-pitched tone and red are indeed perceived to occur simultaneously. If this proportion is zero it means that the high-pitched tone is actually perceived with green.

In the 1.25 Hz condition this proportion is expected to peak above 0.85, since that is the proportion correct matches at 1.9 Hz with 100 % duty cycle in experiment 2 (see Figure 3.3, black bars). For a range of –13.5 to 108 degrees of sound-colour onset asynchrony (see Figure 3.5a) the proportion of the high-pitched tone perceived with red is indeed above 0.85. This range, however, does not show a clear peak at any specific phase shift. Apparently, within a broad range of physical phase differences between

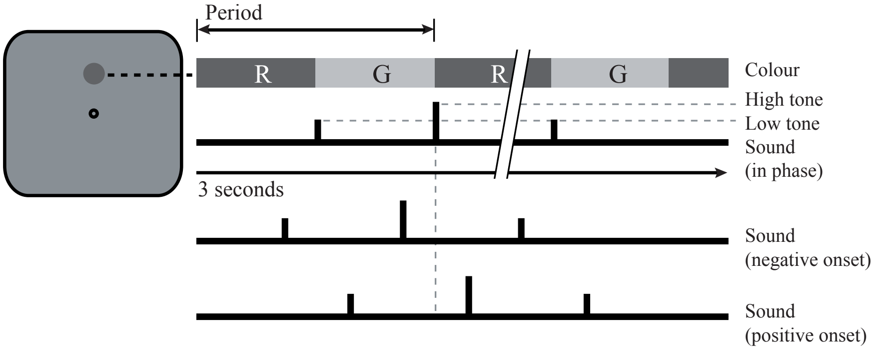

Figure 3.4. Schematic overview of a high-pitched tone co-occurring with red (R) either without sound onset shift (in phase, upper timeline), with a negative sound onset shift (middle timeline) or a positive sound onset shift (lower timeline).

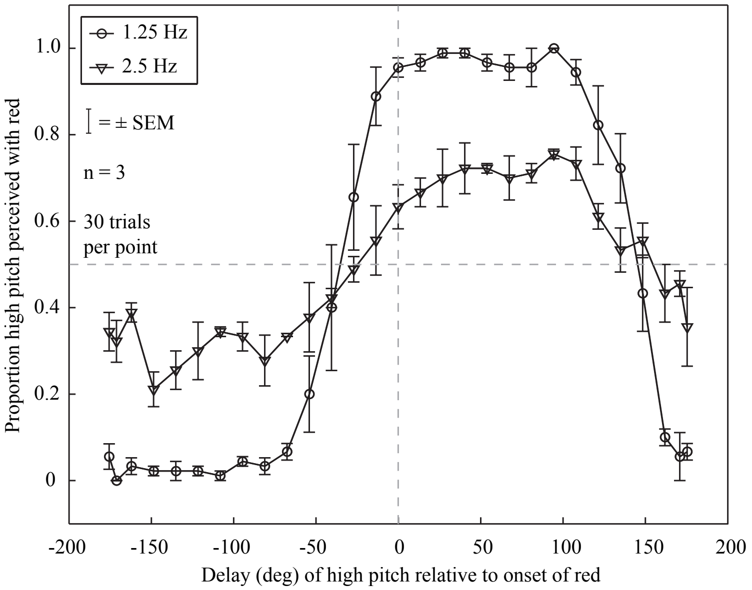

Figure 3.5a. Average proportion of the high-pitched tone being perceived with red as a function of the delay of the high pitch sound relative to onset of red for both frequency conditions in degrees.

sound and colour onset subjects are able to match sound and colour correctly.

For the 2.5 Hz condition, this proportion is expected to peak above 0.67, since that is the proportion correct matches at 3.1 Hz with 100 % duty cycle in experiment 2 (see Figure 3.3, black bars). The maximum proportion is 0.75 in the 2.5 Hz condition at 94.5 degrees phase shift and it remains above 0.67 for a range of 0 to 108 degrees sound-colour onset asynchrony (see Figure 3.5a). Again, a broad range of physical sound and colour onset asynchrony yields similar matching performance. This becomes even more apparent when the data are represented in a polar plot (Figure 3.5b). Rather than showing a clear peak at a specific phase shift, matching performance in both conditions stays the same across a broad range of phasic delay between sound and colour onset. Apparently, even though a (slight) perceptual delay between sound and colour onset might be demonstrated in our setup, it does not appear to play a significant role in matching sound and colour when presented physically synchronous.

When the proportion high-pitched tone perceived in synchrony with red is plotted as a function of the phase-shift between sound and colour, the range of phase-shifts that yield around maximum performance is similar for both alternation rates used: −13.5 to 108 degrees in the 1.25 Hz condition, 0 to 108 degrees in the 2.5 Hz condition (Figure 3.5a). However, when the proportion is plotted as a function of time-delay (in milliseconds; see Figure 3.5c) this range is clearly different: −30 ms to 270 ms in the 1.25 Hz condition, 0 to 135 ms in the 2.5 Hz condition. Interestingly, this means that the range of delays for which sound and colour are perceived as occurring simultaneously depends on the alternation rate. We will return to this in the general discussion.

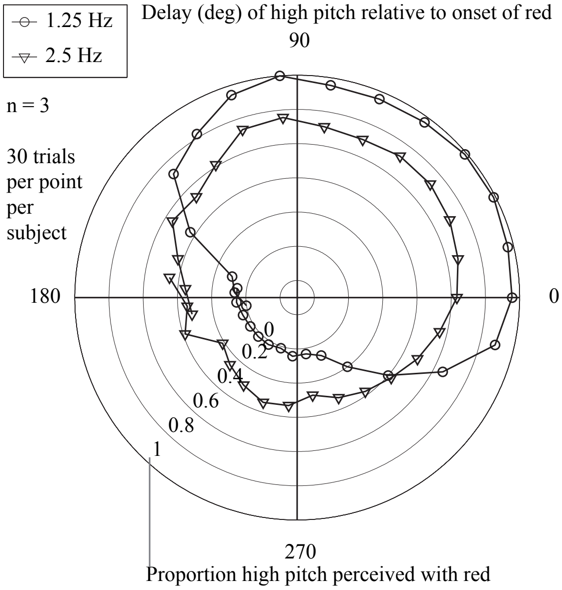

Figure 3.5b. Polar plot of average proportion of the high-pitched tone being perceived with red as a function of the delay of the high pitch sound relative to onset of red for both frequency conditions in degrees.

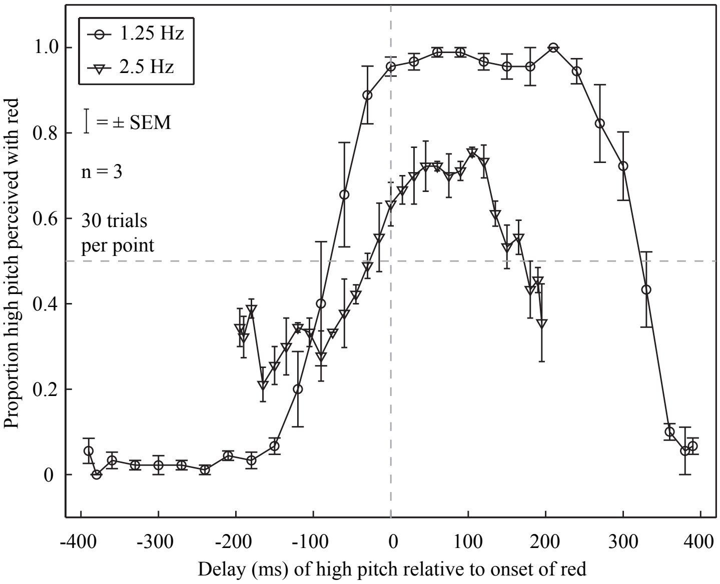

Figure 3.5c. Average proportion of the high-pitched tone being perceived with red as a function of the delay of the high pitch sound relative to onset of red for both frequency conditions in ms.

General discussion

Using an auditory-visual matching paradigm we have found a temporal integration limit for auditory-visual signals around or below 4 Hz. Furthermore, we have shown that this limit is not determined by a temporal integration limit in one of the two sensory systems, which in our experiment demonstrated temporal limits well beyond 4 Hz. This concurs with findings concerning limits of unimodal visual (Holcombe & Cavanagh, 2001) and auditory temporal integration (Viemeister & Wakefield, 1991), although the auditory matching limit found in our experiment is notably lower than Viemeister and Wakefield reported. However, this difference can be explained by the fact that they merely measured detectability, while we measured matching performance.

The temporal limit of integration in the current experiment is not affected by timing of colour offsets, which shows that transient signals created by these offsets are not used for matching sound and colour. When each colour is presented 10 ms performance drops to chance level at presentation rates below the 4 Hz limit. How this very short duration of colour affects matching performance needs further investigation.

We also investigated whether perceptual delay between auditory and visual signals affected the temporal limit we found in experiment 1 and 2. Results from previous auditory-visual studies on perceptual delay have shown that this delay can vary from about 7 to 100 ms (Fujisaki et al., 2004; Keetels & Vroomen, 2005; Lewald & Guski, 2003; Vroomen et al., 2004; Zampini et al., 2005). At two alternation rates below the integration limit found in experiments 1 and 2 (1.25 Hz and 2.5 Hz), matching performance appeared not to be affected by perceptual delay.

Interestingly, the range of auditory-visual onset asynchronies, where matching performance was maximal, was similar at both the 1.25 Hz and 2.5 Hz alternation rate in *phasic delay*, which means that in *time delay* this range scales with alternation rate. Such a frequency dependent scaling of perceptual synchrony has also been demonstrated recently for the perception of natural and artificial auditory-visual motion sequences (Arrighi, Alais & Burr, 2006). Moreover, other reports on auditory-visual perceptual delay show that recalibration takes place after adapting to auditory-visual signals that are presented physically asynchronous (Fujisaki et al., 2004; Vroomen et al., 2004); the point of subjective auditory-visual synchrony due to the perceptual delay shifts with the physical asynchrony. Moreover, other research has shown that even the (simulated) difference between light and sound arrival time can be accounted for as well, when sound and vision (appear to) come from the same spatial location (Sugita & Suzuki, 2004; Alais & Carlile, 2005). Taken together, we conclude that auditory-visual integration has a great capacity for compensating different physical lag times.

Temporal limits of integration

The cross-modal temporal limits presented in the current study are similar to those found for the temporal matching of unimodal visual features, but only when these visual features alternate in time and are spatially separated (Holcombe & Cavanagh, 2001). By separating features spatially, the maximum alternation rate at which colour and orientation can be matched decreases from 18 Hz to 3 Hz. This effect on integration limits of separating visual signals spatially was also confirmed and extended by a recent study by Bartels and Zeki (2006). In addition to spatially separating visual signals, they show a different temporal integration limit for matching visual signals *within* one visual attribute (3.3 Hz) compared to matching spatially separated signals *across* visual attributes (2.6 Hz). Again, when sensory events differ on an extra dimension (attribute, time and space vs. time and space), the temporal integration limit of this input decreases further. In our experiment, we find a limit of auditory-visual temporal integration that is similar to matching spatially separate visual features, while the sensory input in the current experiment only differs in modality and alternates in time. It could be argued that sound and vision are separated in space in our experiment, since the visual stimuli were presented on a monitor at 86 cm away from the subject, while auditory stimuli were presented through headphones. However, a recent study (Arnold, Johnston, & Nishida, 2005) has found that auditory and visual signals of an event that reach an observer at the same point in time tend to become perceptually bound, even when the input comes from different sources such as a monitor and a set of headphones. Therefore, it is tempting to suggest that adding a modality across which sensory input is integrated has the same effect as adding an extra feature when integrating sensory input within one modality. In other words, sensory modality might be processed as just another feature of sensory events. Recent independent confirmation can be found in a study by Bodelón et al. (2007). They showed that the perceptual temporal resolution is higher for individual features compared to the temporal resolution of perceiving conjunctions of features.

It has been recently proposed (Fujisaki & Nishida, 2005) that attentional or 'mid-level' mechanisms are involved in matching simultaneously occurring sensory events. These mechanisms select the salient features of the sensory input, which are then passed onto a temporal correlator. All selected features that fall into a window spanning about 200 ms become correlated. All correlated input is integrated into one multi-sensory percept. This results in a temporal limit of (cross-modal) integration of around 4-5 Hz. Our current results are in agreement with this model, since our limit of around 4 Hz alternation rate is close to that of the model. Thus, a similar temporal limit is found when either a matching paradigm or a synchrony/asynchrony judgment paradigm is used (e.g. Fujisaki and Nishida, 2005). Apparently, the temporal limit of auditory-visual (a)synchrony sensitivity is similar to the temporal limit of auditory-visual integration. Note however, that Fujisaki and Nishida have constructed this model of auditory-visual integration without taking Holcombe and Cavanagh's result (of a higher unimodal matching limit when visual features are spatially superimposed) into account. To incorporate this (and our) results into their model, their temporal correlation window should vary in size, depending on the number of features (including modality) to be matched.

Both in our paradigm and in synchrony judgments, *integration* of sensory inputs will enable observers to make a judgment about the events presented to them (Fujisaki et al., 2006; Fujisaki & Nishida, 2005). In comparison, auditory-visual temporal-order judgments, which are also used to determine auditory-visual sensitivity, rely heavily on *segregation* of the auditory-visual signals (Keetels & Vroomen, 2005; Vroomen et al., 2004). Results for this kind of experiment typically show just noticeable differences (JNDs) in the order of 20-40 ms (which is of a different order of magnitude than the integration results presented here, and by Fujisaki et al. (2005, 2006)). Interestingly, in this type of research, adding a feature on which sensory events differ, specifically spatial location, actually yields better temporal performance. The JND between auditory-visual signals is decreased when sources of auditory and visual input are separated in space (Keetels & Vroomen, 2005; Zampini, Shore & Spence, 2003). Shore, Spence and Klein (2001) did not find such a performance benefit of spatial location in unimodal visual temporal order judgements (although their experiments were designed to answer different questions). Future research should thus focus on investigating whether adding a spatial dimension to auditory-visual matching would yield an increase in auditory-visual matching performance (comparable to auditory-visual segregation experiments) or a drop in performance due to the addition of yet another feature (comparable to unimodal matching/integration experiments, e.g. Holcombe & Cavanagh, 2001).

Chapter 4

Expect the unexpected: top-down suppression of attentional capture in time

Abstract

The ongoing debate whether attentional capture is stimulus driven or top-down influenced has recently focused on the temporal properties of task-irrelevant events leading to attentional capture. Two recent studies have proposed that attentional capture occurs when the task-irrelevant event is either temporally unique (Von Mühlenen, et al. 2005) or when its occurrence is temporally unexpected (Lamy, 2005). In the latter case expected task-irrelevant events are top-down suppressed, while unexpected events are not.

Here we show that attentional capture by task-irrelevant events depends on the timing of task-relevant events (target onset). Before task-relevant events occur, task-irrelevant events can be ignored. But, once the task-relevant event has occurred, task-irrelevant events do influence our percept. Task behaviour can be split in a *waiting mode* and a *searching mode* with different susceptibility to attentional capture. This suggests that attentional capture is not purely bottom-up, but is mediated by top-down processes as well.

Introduction

It is still a matter of debate whether salient objects or events are able to capture attention automatically or whether these objects and events can be suppressed in a top-down manner (e.g. Folk, Remington, & Johnston, 1992; Leber & Egeth, 2006; Theeuwes, 2004; Yantis & Egeth, 1999; Yantis & Hillstrom, 1994). A recent review suggests that attentional capture is a bottom-up phenomenon by default, but it can be modulated in a top-down manner (Ruz & Lupianez, 2002). A problem for pure bottom-up explanations of capture is that not all objects and events automatically capture attention. For example, Theeuwes (1995) reported that an abrupt luminance change captures attention, whereas a colour change does not. This has led to a couple possible explanations for attentional capture such as the new-object-hypothesis (Yantis & Hillstrom, 1994; Yantis & Jonides, 1984) or a delayed-signal-hypothesis (Posner, Snyder, & Davidson, 1980). In addition to these explanations results in a recent article by Von Mühlenen, Rempel, and Enns (2005) suggest that in order to capture attention, an event has to be unique both spatially and temporally. This supports the new-object-hypothesis: if an event is both spatially and temporally unique it could be seen as a new event/object. In the Von Mühlenen et al. (2005) study, it was shown that a task-irrelevant event (a task-irrelevant colour or motion change) only yielded attentional capture when it did not coincide with the onset of either the display or the target. This led them to conclude that a task-irrelevant event must be temporally unique as well to capture attention. Focusing on a thus far relatively little examined temporal factor has indeed given more insight and shows that spatial saliency is not the sole predictor of attentional capture.

Another recent paper on the temporal properties of task-irrelevant events has shown that even temporal uniqueness might not be sufficient to yield attentional capture (Lamy, 2005). She showed that capture is modulated by an observer's expectations regarding the temporal properties of a task-irrelevant event: when its occurrence is expected, be it temporally unique or not, no capture is reported. This expected time of occurrence for an observer is the average time of occurrence when timing is varied randomly. Attentional capture is only evident when a task-irrelevant event occurs at a (for the observer) unexpected moment. This implies a role for top-down influence in attentional capture.

However, based on the experiment by Von Mühlenen et al. (2005), both a temporal uniqueness as well as a temporal expectation explanation (cf. Lamy, 2005) for capture are still possible. The range of onset asynchronies (SOAs) between the task-irrelevant event (the uninformative colour or motion cue) and the target in their experiment was either -1000, -150, 0 or 150 ms. Since a 1000 ms interval between cue and target onset is too large to yield any capture effect, the critical range becomes -150 to 150 ms. (If anything, 1000 ms SOA might yield inhibition of return, since the SOA is larger than 250 ms (Theeuwes & Chen, 2005)). In this range, absence of capture at 0 ms SOA can be explained by the account of both researchers. At 0 ms SOA between cue and target, the cue is not temporally unique, but it also has the expected timing.

Here, we further investigate to what extent the temporal properties of the task-irrelevant and target events can influence attentional processing by adapting the paradigm by Von Mühlenen, using a different range of cue and target onset asynchronies (-150 to 50 ms). More specifically, we set out to investigate whether we can find attentional capture at 0 ms SOA, since temporal uniqueness and no capture at 0 ms SOA is inferred from slope ratios based on set sizes of 3 to 7 elements in a stimulus display. Slope ratios would be a good measure if they were based on trials in which attention was equally deployed. However, 3 is a quite small set size to assume the stimulus is processed the same way as a 7 element-stimulus. Three items falls within the subitising range and can thus be processed 'at a glance'/all at once. In other words participants process the 3 element stimulus pre-attentively. Thus, slope ratios are inferred from data points that might be the result of different attentional processes. It would therefore be better to examine 'raw' RT differences, which makes it also easier to compare to Lamy (2005).

The reaction time (RT) to target will be shorter when a target is cued by the task-irrelevant event compared to situations where a non-target is cued by this event, as was the case in Von Mühlenen et al. (2005) and Lamy (2005). In other words, the reaction time differences will be positive, when capture occurs. In our experiments the range of -150 to 50 ms cue-target SOA would therefore yield no capture at 0 ms according to a temporal uniqueness account (Von Mühlenen et al. (2005)), since the cue is not temporally unique in relation to target onset. In contrast, if observers build up an expectancy of the timing of the colour change, this would yield no capture at -50 ms cue-target SOA according to Lamy (2005, see also the expected RT behaviour in Figure 4.1).

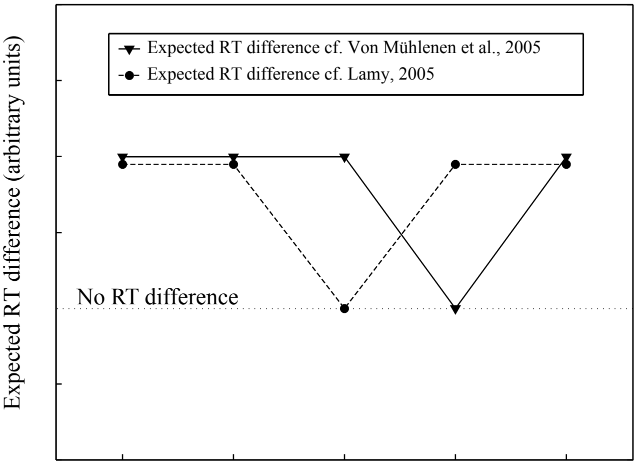

Figure 4.1. Expected RT difference (in arbitrary units, invalid minus valid cue conditions) according to either a temporal uniqueness account of attentional capture (no capture at 0 ms SOA, cue is not temporally unique event cf. Von Mühlenen et al., 2005) or a temporal expectancy account of capture (no capture at -50 ms SOA, cue occurs at the expected time cf. Lamy, 2005).

We show that a task-irrelevant event can successfully be suppressed before target onset, and that an attentional capture effect is present only *after* target onset. This suggests top-down suppression of task-irrelevant events as long as target events have not yet occurred. In a separate experiment, we varied the target onset while keeping the timing of the task-irrelevant event constant (with the same range of SOAs) and find identical results: a task-irrelevant event only captures attention if it occurs after target onset. Target onset thus divides behaviour in this type of capture paradigm in two separate modes: waiting and searching. The latter seems to be susceptible to attentional capture.

Experiment 1

Methods

Participants

Nine observers (7 naïve as to the purpose of the experiment) voluntarily participated in the experiment. All had normal or corrected to normal vision. Two observers were authors (JB and HH).

Stimuli

Stimuli were created and presented using Matlab® 5.2.1 and the Psychophysics Toolbox extensions (Brainard, 1997; Pelli, 1997) on a Macintosh G4 computer. Stimuli were presented on a 19 inch Iiyama Pro Vision Master 454 monitor set to a resolution of 640 by 480 pixels and a vertical refresh rate of 200 Hz yielding 5 ms timing precision.

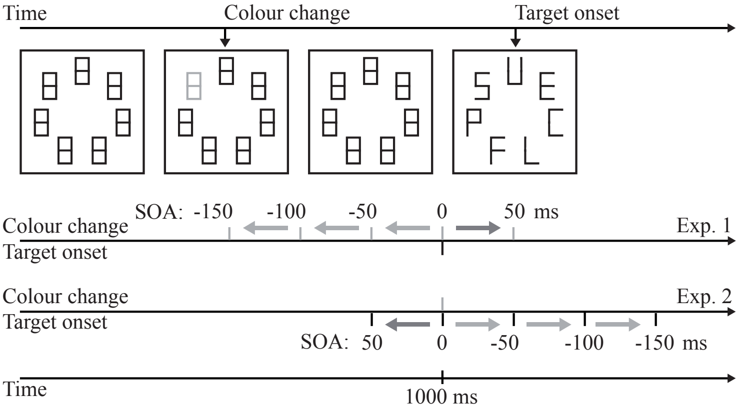

Figure 4.2. Schematic representation of the stimulus. The displays from left to right indicate the stimulus sequence in which a colour change (indicated by the gray element) and target onset formed the manipulated events. Time lines indicate when a 50 ms. colour change was presented in one the elements (gray spikes) relative to target onset (black spikes) i.e., indicate cue-target SOA. Among these letters a target letter was present, in this case an U.

Participants were placed in a head and chin rest such that the distance to the screen was 57 cm.

A circular array of seven evenly spaced digital eights was presented for 1 second (see Figure 4.2). After this interval, between two and four line segments were removed from these eights, changing the display in a circular array of seven letters. These letters were E, P, S, C, F, L and either a letter H or U. The letters and the eights were grey on a black background (luminance: 10.7 cd/m^2 and 0.26 cd/m^2, respectively), had a height of 1.94 degrees, and a width of 1.22 degrees visual angle. The elements were arranged on an imaginary circle with a radius of 5.5 degrees visual angle with a white fixation point (luminance: 133 cd/m^2) with diameter 4.2 arcminutes of visual angle at the center.

The task-irrelevant event that was used in this capture paradigm was an irrelevant colour change in one of the seven elements in the display. One element changed from grey to red (luminance: 35.8 cd/m^2) for a duration of 50 ms. In other words, an auto-cue (an event occurring in one the elements itself instead of around one of the elements; an allo-cue, see Rauschenberger, 2003) was used as the task-irrelevant event. Note that, since red is not made equiluminant to grey of the others elements, this colour change implies a luminance change as well, making this change a particularly salient event. This colour cue occurred at different moments during the presentation of the stimulus: either at 150, 100, 50 ms before, simultaneous with or 50 ms after target onset. Each element was cued 24 times, for a total of 840 trials (5 timing conditions x 7 elements x 24 repetitions). In 120 trials the target letter was cued (valid cue), and in 720 a non-target letter was cued (invalid cue). The probability of the target being cued was therefore 14%.

Task and procedure

Observers made a speeded 2-alternative forced choice (2AFC) decision whether the array contained a letter H or U. All letters remained visible until the observer responded. Immediately after a response was given the display was cleared and the next trial started 200 ms after that. Note that, the latest a colour change could start is 50 ms after target onset in this experiment, average response time was around 600 ms after target onset, and the earliest the next colour change can occur was 850 ms after display onset. Thus, the shortest time between two colour changes was (600 - 50 + 200 + 850) = 1600 ms. It therefore seems highly unlikely that transient activity from one colour change in one trial will interfere with the colour change in the next trial. Observers were instructed to respond as quickly as possible, but to keep errors at a minimum as well. After ten practice trials, observers carried out all 840 trials in one block, taking about 30 minutes per observer.

Results and Discussion

Mean reaction times (RTs) to target in valid (a colour change in the target element) and invalid (a colour change in one of the six distractors) conditions were first calculated for each cue-target onset asynchrony (each SOA). An RT difference was then calculated for each SOA by subtracting the mean RT in valid trials from the mean RT in invalid

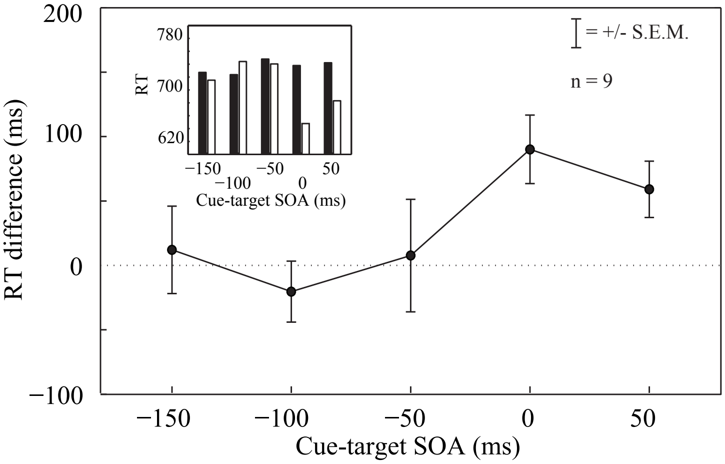

Figure 4.3. Average RT difference (RT to target in invalid minus valid cue conditions) in ms as a function of distractor target SOA from experiment 1. Inset: average RT on invalid (black bars) and valid (white bars) trials.

trials. A positive difference therefore implies a performance benefit from the colour cue: observers are faster to report the target letter when the colour change occurred in that element. Accordingly, a negative RT difference indicates a performance decrement. Both performance costs and benefits can be described as manifestations of attentional capture (see Rauschenberger, 2003).

The RT differences for all nine observers of experiment 1 are averaged and plotted in Figure 4.3.

There is a clear performance benefit (a positive RT difference) from valid cueing at 0 and 50 ms cue-target SOA. Repeated measures ANOVA confirmed this effect of SOA on RT difference ($F_{(4,32)} = 9.368$, $p = 0.002$). Preceding target onset, there is no evidence of attentional capture: the average RT difference does not significantly differ from zero RT difference (indicated by the dotted line in Figure 4.3) in this experiment (running t-tests for these SOAs showed that $t < 1.08$ and $p > 0.34$ for cue-target SOA –150, –100 and –50 ms, while $t > 2.70$ and $p < 0.05$ for cue-target SOA 0 and 50 ms). The clear performance benefit, at and after target onset, points towards an explanation for attentional capture that does not include temporal uniqueness (Von Mühlenen et al., 2005). It could be argued that for the time points preceding target onset, the transient activity induced by the colour change has faded. However, that would mean that this activity by the colour change at cue-target SOA –50 ms (duration of 50 ms and starting 50 ms before target onset, thus ending at target onset) fades within the duration of the colour change and thus does not yield RT differences. This seems quite fast. It would suggest a very rapid recovery from capture, faster than reported by Theeuwes et al. (2000).

Alternatively, the absence of capture preceding target onset might be attributed to target onset timing being fixed. In experiment 1, target onset was always 1000 ms after display onset. Observers could therefore adapt a strategy in which they just ignore the first second of each trial. To control for this possible strategy, we ran experiment 2 as a control.

Experiment 2

As mentioned above, in the first experiment, target onset was always 1000 ms after display onset. Observers could therefore adapt a strategy in which they just ignore the first second of each trial. In comparison to the first experiment the timing of target onset was therefore varied to lose this possible predictability of target onset (see Figure 4.2).

Methods

Participants

Five observers (4 naïve as to the purpose of the experiment) voluntarily participated in the experiment. All had normal or corrected to normal vision. One observer was author (JB). All observers in this experiment had participated in experiment 1 as well.

Stimuli and procedure

Stimuli, procedure and target-distractor SOAs were identical to those in experiment 1. Timing of target onset was varied relative to a fixed cue onset (1000 ms after display onset, see Figure 4.2). Participants ran the 840 trials in two blocks of 15 minutes.

Results

Mean reaction times (RTs) to target in valid (a colour change in the target element) and invalid (a colour change in one of the six distractors) conditions were again calculated first for each cue-target SOA. In the same way as in experiment 1, RT difference was then calculated for each SOA by subtracting the mean RT in valid trials from the mean RT in invalid trials. These RT differences for all five observers are averaged and plotted in Figure 4.4.

Again, it can be seen that before target onset, even if its timing is unpredictable in this experiment, task-irrelevant events are suppressed. RT difference is positive only when the cue occurs simultaneous with or after target onset. Repeated measures ANOVA confirmed again the effect of SOA on RT difference ($F_{(4,16)} = 8.652$, $p < 0.001$). Preceding target onset, there is no attentional capture: average RT difference does not significantly differ from zero RT difference (indicated by the dotted line in Figure 4.4) in this experiment (running t-tests for these SOAs showed that $t < 1.72$ and $p > 0.16$ for cue-target SOA −150, −100 and −50 ms, while $t > 1.93$ and $p < 0.05$ for cue-target SOA 0 and 50 ms).

With this second experiment we replicated the results from experiment 1 and ruled out a possible strategy of ignoring the first second of each trial in experiment 1.

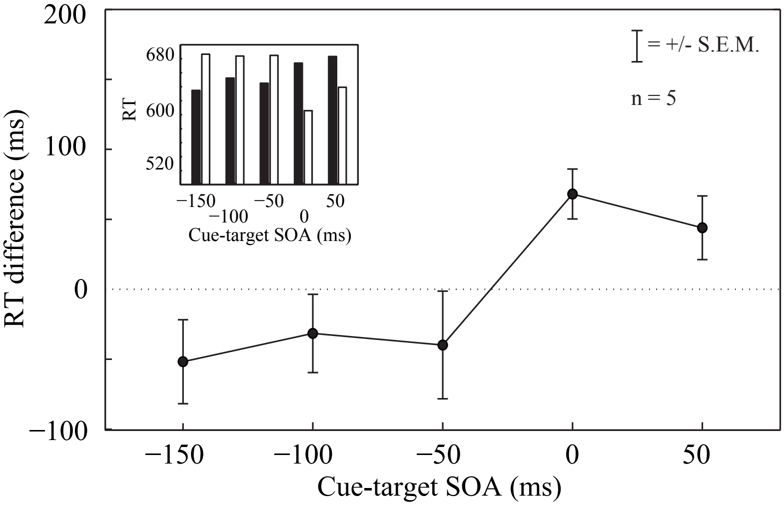

Figure 4.4. Average RT difference (RT to target in invalid minus valid cue conditions) in ms as a function of distractor target SOA from experiment 2. Inset: average RT on valid (white bars) and invalid (black bars) trials.

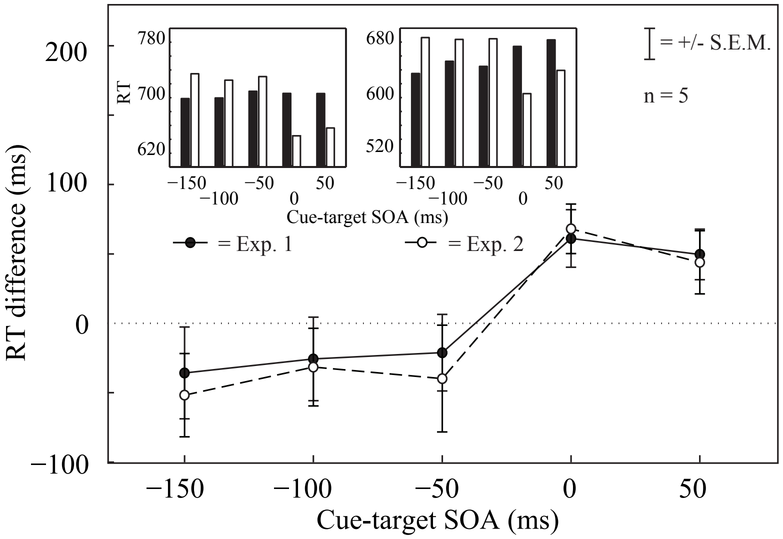

Figure 4.5. Average RT difference (RT to target in invalid minus valid cue conditions) in ms as a function of distractor target SOA from the five observers that participated in the two experiments (solid line, closed symbols = experiment 1; dashed line, open symbols = experiment 2). Insets: average RT on valid (white bars) and invalid (black bars) trials. in experiment 1 (left) and experiment 2 (right).

Moreover, repeated measures analysis of the data from the five subjects that participated in both experiments (plotted in Figure 4.5) with experiment (timing of target onset fixed versus variable) as between-subjects factor and cue-target SOA as within-subject factor confirmed the effect of timing ($F_{(1,4)} = 22.301$, $p < 0.001$) and the fact that timing of target onset (experiment 1 versus experiment 2) did not yield different results ($F_{(1,4)} = 0.012$, $p = 0.809$). When the colour change occurs simultaneous with or after target onset in the target letter, independent of the timing of target onset itself, observers report the target letter faster compared to when this colour change occurs in a non-target letter at or after target onset.

Discussion

The debate whether attentional capture is stimulus-driven or top-down modulated has mainly focused on spatial and other non-temporal factors (e.g. Folk & Remington, 1998; Folk, Remington, & Johnston, 1992; Folk, Remington, & Wright, 1994; Lamy, 2005; Lamy & Egeth, 2003; Lamy, Leber, & Egeth, 2004; Leber & Egeth, 2006; Theeuwes, 1990; Theeuwes, 2004; Yantis & Egeth, 1999; Yantis & Hillstrom, 1994). Recently, however, two studies have looked into the temporal properties of task-irrelevant events as well. These studies suggest different explanations for attentional capture. Lamy (2005) proposes that irrelevant events capture attention when they occur at an unexpected point in time, whereas Von Mühlenen, Rempel, and Enns (2005) argue that an event needs to be temporally unique to induce attentional capture. The latter assumes that every event that is temporally unique captures attention, whereas the former shows this to be the case only when the timing of this event is unexpected.

Here, we show, using an adapted version of Von Mühlenen et al.'s stimulus (2005), that temporal uniqueness alone is not sufficient to cause attentional capture. We show that observers are quite capable of suppressing task-irrelevant events. An irrelevant colour change (50 ms duration) occurring in one of the elements in a seven element cueing paradigm does not affect reaction time (RT) to the target element when target onset has not yet occurred (up to 150 ms before target onset). Conversely, a performance benefit is observed when the irrelevant event occurs in the target element simultaneously with or after target onset. After, or simultaneous with target onset the RT is shorter when the irrelevant event validly cued the target. Note that the rapid colour change is a dynamic cue. As argued by Rauschenberger (2003) this will yield attentional capture, while a static differently coloured item in these experiments would more likely result in perceptual guidance. Thus, using this cue it is shown that observers are well able to withhold making a shift of attention to the element that underwent an irrelevant change, when the relevant target onset has not yet occurred. Even when target onset is made unpredictable, as in Experiment 2, no capture effects were present before target onset. Note that RT difference seems to show a negative trend preceding target onset. However, this might be due to the smaller number of participants in Experiment 2. Having more participants in Experiment 1, next to the five that participated in both experiments, results in this negative RT difference before target onset being absent, while the positive RT difference still is present.

The current findings seem to be in conflict with the notion that all temporally unique events cause attentional capture in a bottom-up fashion, as proposed by Von Mühlenen et al. (2005). They show that, when target-distractor SOA is 0 ms, the RT slope (ms/per item more in the display) is the same for valid trials and for invalid conditions. RT slope is their measure for attentional capture and is used to conclude that only temporally unique events cause capture. However, a closer inspection of their data shows that there are, in fact, RT differences at temporally non-unique occurrences of the distracting task-irrelevant event. Despite having equal RT slopes, there are clear RT differences between valid and invalid conditions at SOAs where the colour change is not temporally unique: in a seven element display, a valid cue at 0 ms SOA yields a RT to target that is 89 ms shorter than an invalid cue at the same SOA. In other words, though the RT *slope* may be the same, there still is a considerable *offset* difference between valid and invalid cueing conditions (see also Figure 4.6 for the RT differences from Von Mühlenen et al. (2005) in comparison to the current experiments). Moreover, as mentioned in the introduction, the number (3, 5 and 7) of items that have to be searched for a target seem to be somewhat small to infer an ms/per item search slope from. Especially when detecting a target in a 3 element display, this is a number in the subitising range and yields instantaneous pre-attentive parallel processing instead of serially searching for a target in a seven element display. Thus, slope ratios are inferred from data points that might be the result of different attentional processes.

Our results can be considered as an extension of the top-down modulated account for attentional capture by Lamy (2005). We also find that observers can suppress responding to a task-irrelevant event. However, it is not the expected timing of this event

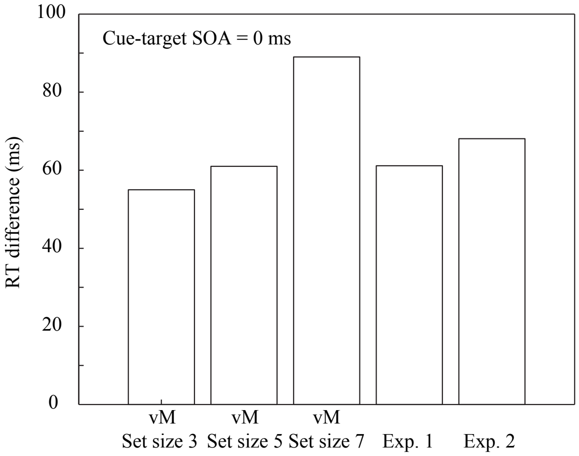

Figure 4.6. Data adapted from Von Mühlenen et al. (2005) and data from the current experiments at cue-target SOA of 0 ms. Data are (re)plotted here in terms of RT difference (RT to target in invalid minus valid cue conditions, in terms of Von Mühlenen et al. (2005), 'no change' minus 'change' conditions). RT difference is plotted for each set size from Von Mühlenen's experiment separately (left three 'vM' bars).

per se that determines this suppression. Observers can withhold a shift of attention when target onset has not occurred yet.

Note that in Lamy (2005) there is an RT difference of cueing targets and non-targets before target onset. However, the RT differences in these studies are small (20/30 ms difference) compared to the current findings (around 60 ms) and are obtained by using allo-cues (cues that do not spatially overlap with the elements, see Rauschenberger, 2003). It is possible that these allo-cues are more salient cues than the auto-cues (colour change) in the current experiment. According to Folk and Remington (1998), top-down control of attentional capture by salient cues might only hold up to certain levels of salience of cues. They argue that a distractor can be of such extreme intensity that it disrupts all ongoing processing. This is confirmed by experiment 3 in Lamy (2005), where a capture effect (an RT difference) is present at all distractor-target SOAs when salience of a diamond shaped allo-cue is increased. A top-down modulation of attentional capture is no longer possible in those conditions. In our experiments we find that the cue, while maintaining the same saliency across SOAs, only affects RT when target onset has occurred.

Searching versus Waiting

It appears as if observers adapt a *waiting mode* until target onset, after which they enter a *searching mode* that is susceptible to cueing (expressed here as a performance benefit when the cue is valid). Together the two experiments clearly show a task-dependent top-down modulation of attentional capture: when task-relevant events have not occurred and observers are in waiting mode, task-irrelevant events are suppressed. Interestingly, researchers have proposed an initial bottom-up sweep from which recovery takes around 50-100 ms (Theeuwes et al., 2000). This bottom-up seep is supposed to not be susceptible to top-down influences and turns into inhibition of return after 250 ms SOA (Theeuwes & Chen, 2005). In our experiments this should have resulted in bottom-up cueing effects across the entire range of SOAs used, since it only spans 200 ms. For example, if recovery from capture is 100 ms we should certainly have found RT differences when colour change preceded target onset by 50 ms. Our experiments indicate that this initial bottom-up sweep is either shorter than thus far reported or that it can be top-down suppressed as long as task-relevant events have not yet occurred. As such, the current experiments point towards a view of initial attentional capture that is not purely bottom-up, but might be influenced by early top-down processes as well.

Chapter 5

Search time critically depends on irrelevant subset size in visual search

Abstract

Our visual system can not fully process all of its input (Wolfe, 1994). To solve this problem some of this input is discarded, while other parts of this input are selectively processed. From visual search literature it is unclear how well one set of items can be selected that differs in only one feature from target (a 1D set), while another set of items can be ignored that differs in two features from target (a 2D set).

We systematically varied the percentage of 2D non-targets to determine the contribution of these non-targets to search behaviour. Increasing the percentage 2D non-targets, that have to be ignored, is expected to result in increasingly faster search, since it decreases the size of 1D set that has to be searched. Observers searched large displays for a target in the 1D set with a variable percentage of 2D non-targets.

Interestingly, when the search displays contain 5% 2D non-targets, search time is longer compared to search time in other conditions. This effect of 2D non-targets on performance is independent of set size. Inspection of the saccades revealed that saccade target selection does not contribute to the longer search times in displays with 5% 2D non-targets. Occurrence of longer search times in displays containing 5% 2D non-targets might be attributed to covert processes related to visual analysis of the fixated part of the display. Apparently, visual search performance critically depends on the percentage of irrelevant 2D non-targets.

Introduction

Typically when perceiving the world, we need to select those parts of the world that are of interest to us. We need to make this selection since our visual system has a limited capacity and can not process all aspects of the visual world at once. Visual search paradigms are a tool to measure how we select this information.

A hallmark result of this field of research is that visual search can be divided into two types of search (Neisser, 1963; Egeth & Dagenbach, 1991): parallel and serial search (Foster & Ward, 1991; Treisman & Gelade, 1980; Treisman & Gormican, 1988). Search-time slopes are generally acknowledged to be a useful diagnostic tool to make a distinction between parallel and serial search. A well-established explanation for different search-time slopes between parallel and serial search is that they differentially draw on attentional resources.

Serial visual search can be attributed to different factors. It can be due to a small difference between target and non-target (Duncan & Humphreys, 1989; Pashler, 1987). Another possibility is that there are two types of distractors, each sharing a different feature with the target, making it hard to find the right combination of features that defines the target. In this so-called *conjunction search*, each item has to be compared to other non-target combinations of features, making search relatively slow (Treisman & Gormican, 1988; Treisman & Sato, 1990). However, studies using conjunction search stimuli (Bacon & Egeth, 1997; Shen, Reingold & Pomplun, 2003; Zohary & Hochstein, 1989) have shown that search time decreases, if one of the two types of distractor conjunctions is less frequent than the other, while the total set size remains constant. This results in a quadratic relationship between set size of one type of distractor and search time; when one set of distractors is either very frequent or very infrequent, search time is shorter than when both distractors types are equally frequent. This quadratic set size-search time relationship in conjunction search is thought to be a bottom-up effect of the smaller more salient subset. However, Bacon and Egeth (1997) show that this quadratic relationship between type of distractor and search time can be top-down attentionally guided. They show that in conjunction search, response time can be predicted by a linear top-down relationship between set size of one type of distractor and search time on the one hand and a bottom-up quadratic relationship between that same set size and search time on the other hand.

Now consider feature search again, where a target is defined by a unique feature. Suppose there are two subsets of distractors in the search display with a constant total set size, where one subset can never contain the target, since it differs on more than the target feature alone (e.g. an irrelevant subset). The prediction would be that if the relevant subset can be successfully selected, search time should show a decrease as the irrelevant subset size is increased. Williams (1966), for example, has shown that search time decreases, when the target is coded based on one specific feature (e.g. colour or shape). When observers know what feature defines the search target, search time for that target decreases. Observers appear to use colour to select a subset of items in a search display. Note that being able to select a subset of items (Luria & Strauss, 1975) or not (Zelinsky, 1996), again might depend on target-distractor similarity (Hooge &

Erkelens, 1999). Alternatively, if there is still some bottom-up activation due to a smaller subset being more salient, we expect a combined quadratic and linear relationship (as in Bacon and Egeth, 1997) between irrelevant subset size and search time.

Such a feature search task as described above can be found in a unpublished observation by Wertheim (1981), where stimuli were used that consisted of slides of photographed stimuli (red and white thumbtacks on a blackboard, with one smaller white thumbtack as target). Search time data from this experiment, however, could not be described by a linear or the combination of a quadratic and linear relationship between irrelevant subset size and search time. Interestingly, there appeared to be a search time increasing effect on top of a quadratic/linear pattern (cf. Bacon & Egeth, 1997) when the irrelevant distractors were infrequent (around 5% of the total set size). This can be seen in Figure 5.1, where a quadratic/linear pattern – like Bacon and Egeth (1997) have used to model search time – is least square fitted to the data. Around 5% irrelevant subset size, the data deviate to a large extent from the function describing the best quadratic/linear fit on the data (the mean of absolute residuals is 517 ms for percentages between 0 and 5% (SD = 415 ms), whereas absolute residuals at higher percentages have a mean of 181 ms (SD = 197 ms), thus showing a better fit).

However, since the data of Wertheim's study were acquired using slides of photographed thumbtacks, the deviation from the predicted relationship between irrelevant subset size and search time around 5% irrelevant subset size, could be due to luminance differences between the different stimuli, rendering the irrelevant non-targets

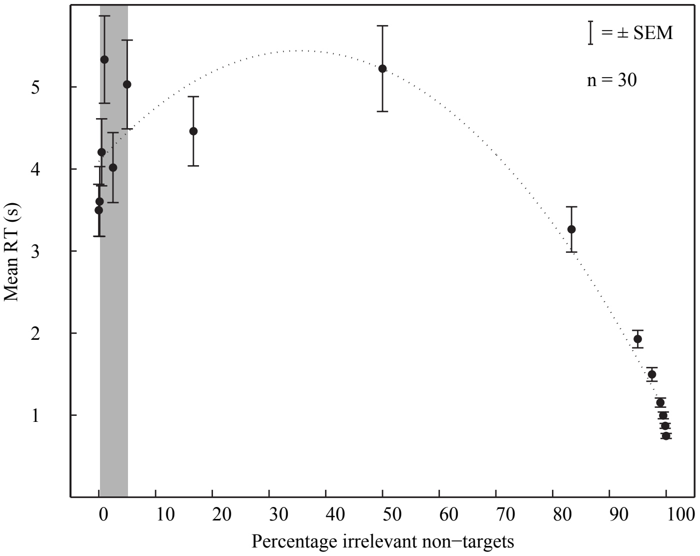

Figure 5.1: results from Wertheim (1981). Search time (RT in seconds) plotted as a function of the percentage of irrelevant non-targets. The dotted line indicates the best least squares fit of a quadratic/linear pattern as in Bacon and Egeth (1997). The gray area indicates 0-5% 2D non-targets.

perhaps more salient and thus causing an extra saliency effect around 5% irrelevant subset size. Hence, we try to replicate this study here using a computer setup to test whether we can find the same results under more controlled conditions.

Experiment 1

We examined whether increasing the size of an irrelevant subset, while keeping total set size constant, would decrease search time due to successful selection of the colour coded relevant subset. Alternatively, we would show the same results as the experiment by Wertheim (1981). Next to the relevant subset being colour coded, a relative large similarity between the relevant non-targets and the target in that relevant subset should make subset selection more efficient and decrease search time as the irrelevant subset is increased in size, because the relevant subset is thus concurrently decreased. The irrelevant subset of items differed from target on two dimensions (colour and size) while the relevant subset only differs in size from target. We therefore refer to the irrelevant subset of items as 2D non-targets and to the relevant as 1D non-targets.

Methods

Participants

Four observers (3 naïve as to the purpose of the experiment) voluntarily participated in the experiment. All had normal or corrected to normal vision.

Stimuli

Stimuli were created and presented using Matlab® 5.2.1 and the Psychophysics Toolbox extensions (Brainard, 1997; Pelli, 1997) on a Macintosh G4 computer. Stimuli were presented on a 19 inch Iiyama Pro Vision Master 454 monitor set to a resolution of 1280 by 1024 pixels and a vertical refresh rate of 85 Hz. Participants were placed in a head and chin rest such that the distance to the screen was 57 cm.

Each trial consisted of a search display. This display consisted of a rectangular array (height/width ratio 2:3) containing 600 small (0.71 degrees of visual angle) grey dots (luminance: 15.1 cd/m2), separated horizontally and vertically 0.5 degrees visual angle on a black background (luminance: 0.26 cd/m2). All positions of the dots were randomly displaced 0.24 degrees of visual angle horizontally and vertically. Each search display contained a target, which was a smaller grey dot (0.42 degrees of visual angle). The target thus only differs from all elements in one dimension, namely size.

In addition to this size difference in size compared to target, in 16 out of 17 conditions a given number of non-targets dots in the display were replaced by dots that differed in two dimensions from the target. Apart from being bigger than the target dot (0.71 degrees visual angle), these dots were red. This colour was made equiluminant to the grey of the other dots using a flicker photometry procedure (Cavanagh, Tyler, & Favreau, 1984). The number of red 2D non-targets ranged from 0 to 599 in 17 steps (see Figure 5.2 for a search display example). Note that the reverse conditions were tested in pilot experiments as well (red small target amongst different levels of grey 2D distractors).

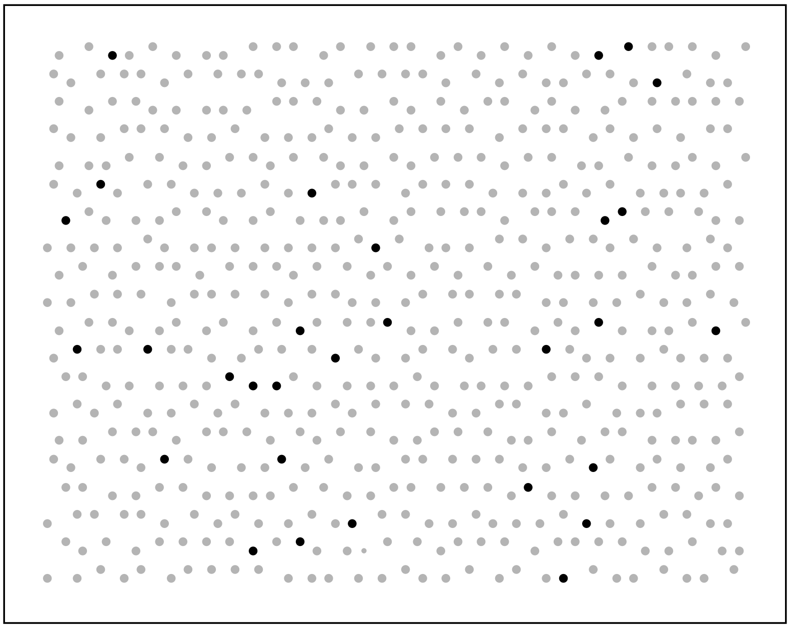

Figure 5.2. Schematic view of a search display. Background of the actual stimulus was black, while 1D non-targets were gray and equiluminant to the red 2D non-targets (here indicated by the black dots). Here 30 out of 600 non-targets form the 2D non-targets set, the smaller gray target dot is present near the low middle of this schematic display.

This showed no difference in performance compared to experiments with grey 1D and red 2D items described here.

Each of these 17 irrelevant subset conditions was repeated 20 times in each subject (340 trials) in two separate blocks containing a random order of 10 repetitions of each irrelevant subset size.

Task and procedure

Each observer first was asked to set their red-grey equiluminant point by repeating the flicker photometry procedure 5 times (Cavanagh et al., 1984), after which the experiment only continued if the standard deviation of these five repetitions of adjusting red to grey was less than 5% of the red colour range.

The task for each observer was to report the target as soon as they detected it. To measure search time we measured the response times of each observers. They were instructed that the target item was always a small grey dot. As soon as observers saw the target they pressed the space bar, after which the search display disappeared and a white rectangular outline appeared. This outline was divided in 4 quadrants, with a dot in one of the quadrants. Observers could move this dot around the quadrants using left and right arrow keys to indicate in which quadrant of the display they found the target. Hitting the escape key ended this indicating procedure. After ending this indicating

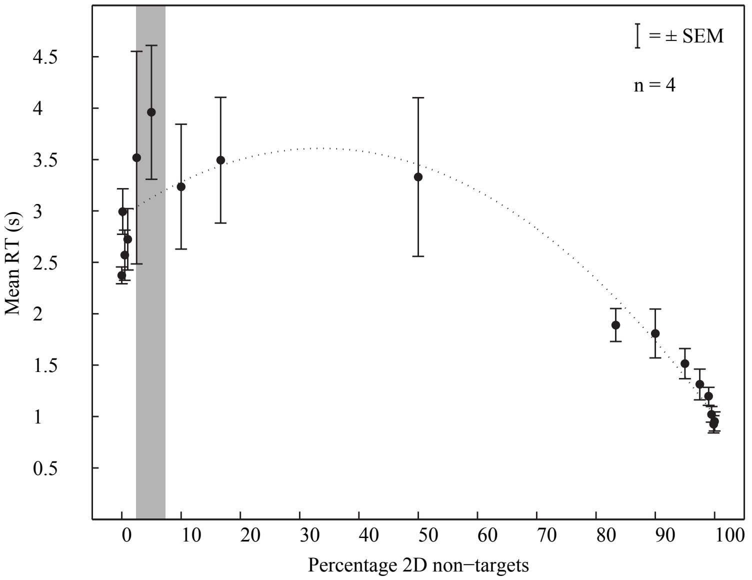

Figure 5.3. Average search time (RT in seconds) as a function of percentage 2D non-targets. The dotted line indicates the best lest squares fit of a quadratic/linear pattern cf. Bacon and Egeth (1997). The gray area indicates 2.5-7.5% 2D non-targets.

procedure, the next trial started 200 ms later. Each observer was given 10 practice trials before the experimental blocks were run.

Results and discussion

The results from experiment 1 corroborated the results found by Wertheim (1981). For each observer the mean response time (RT) in each condition is calculated, after which the averages of these mean RTs are calculated across subjects. The resulting RT curve is plotted in Figure 5.3. Incorrect responses were excluded, which was 2.13% of the data. Contrary to the predicted linear decrease in RT time due to the increasing number of red 2D non-targets (the relevant subset of grey items gets smaller), RT time is again actually longer compared to the condition where there are no 2D non-targets. This increase holds up to 50% 2D non-targets, but has a maximum at 5 % 2D non-targets (30 out of 600).

Results are only comparable to results from Bacon and Egeth (1997) at higher percentages of 2D non-targets. They show that search time in their conjunction search task results can be described by a quadratic bottom-up saliency-based distractor ratio effect, together with a linear decrease in search time due to top-down successful selection of a subset decreasing in size. Fitting our data with this kind of function (a quadratic function centered around 50% 2D non-targets combined with a linear decrease as number of 2D non-targets increases) shows that this can describe the data quite well (see Figure 5.3) at percentages of 10% or higher (mean absolute residual: 99 ms (SD = 63 ms)). What is striking though, is that at lower percentages of 2D non-targets the data can not be fitted well (mean absolute residual: 429 ms (SD = 269 ms)) and are well above the fitted function at 5% (see the grey area of Figure 5.3).

In conclusion, displays with 5% irrelevant items yield longer search times than any other condition. Though the red 2D subset can never contain target in our search displays, a display containing a small number of red dots yields a longer search time, rather than a shorter search time. Search time was expected to be shorter than search time in the 0% red dots condition, since the relevant subset size is smaller. To examine whether this effect on search performance is due to the percentage of 2D non-targets or to an absolute number of 2D non-targets, we tested multiple set sizes in a second experiment.

Experiment 2

To investigate whether search time increases due to a small irrelevant subset of 2D non-targets, or that this is increase is due to an absolute number of these irrelevant non-targets, we repeated experiment 1 in four different set sizes. By investigating search time in different set sizes with four percentages of irrelevant non-targets we could investigate whether the longer search time is set size independent.

Methods

Participants

Four observers (3 naïve as to the purpose of the experiment) voluntarily participated in the experiment. All had normal or corrected to normal vision. None of the participants of this experiment took part in experiment 1.

Stimuli

Stimuli were the same as in experiment 1, except that the display consisted of a rectangular array (height/width ratio 2:3) containing a variable set size. The rectangular array contained either 96, 216, 360 or 600 dots and was centered on the monitor. Again, in some conditions a variable number of grey dots in the display were replaced by red dots. The number of red dots in each trial was manipulated as a percentage of the total number of dots in a display. In the search display either 0, 5, 50 or 95 % of the non-target dots were red.

The resulting 4 x 4 design (number of dots x percentage of red non-targets) was repeated 20 times in each subjects in two separate blocks containing a random order of 10 repetitions per condition.

Task and procedure

The task and procedure for each subject were the same as in experiment 1. The experiment took about 20 minutes.

Results and discussion

Again we find that search times are longest when there are 5% 2D non-targets in the display. The mean search time in each condition for each observer is calculated, after which the averages of these mean search times are calculated across subjects. The resulting mean search time is plotted in panel A of Figure 5.4. Incorrect responses were excluded, which was 2.26% of the data. For each percentage 2D non-targets search time is plotted as a function of the total number of dots in a search display. Immediately apparent from this graph is that for all set sizes search time is longest when there are 5% 2D non-targets in the display. Fitting a linear equation (RT = A * <number of dots> + B) using a least squares method to determine a search time slope for all four percentages of 2D non-targets in a display confirmed this. The slope (parameter A) for the 5% red dots conditions is twice as steep (4.71) as in the 0% (2.34) condition, 1.5 times as steeps as in the 50% condition (3.19), and about three times as steep as the slope in the 95% red dots condition (1.50, see Figure 5.4, panel B). Each of the R^2s of the fitted functions is above 0.94.

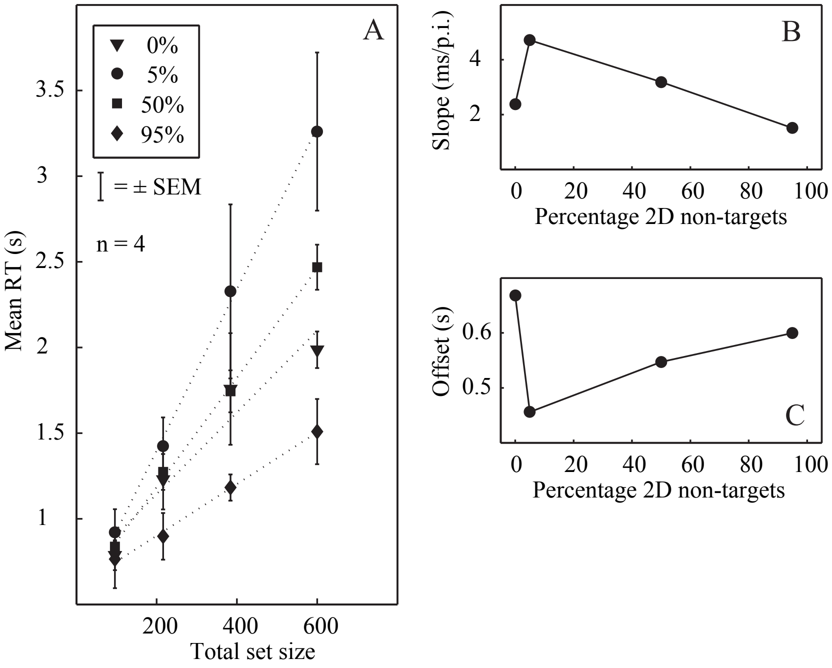

Figure 5.4. Average search time (RT in seconds) for each of the four percentage conditions (panel A). The results of a least squares fit of a linear function through the data are plotted as dotted lines. The parameters slope and offset of this fit are plotted as a function of percentage 2D non-targets in the right panels (slope in panel B and offset in panel C).

Thus, the striking result again is that with a small number of 2D non-targets in a display search times are actually longer than in the other conditions, and do not depend on an absolute number of 2D non-targets. Independent of total set size, 5% irrelevant non-targets yield the longest search time. Though red dots can never be the target in a search display, a small number of red dots lengthens search time, rather than shorten it.

These longer search times can be attributed to different underlying processes. During search in large displays, observers usually make eye movements. To get more insight in what may make search slow with 5% 2D non-targets, we measured eye movements of five observers performing the same task as in experiment 1.

Experiment 3

To investigate whether we can attribute longer search times in the 5% 2D non-target condition to different underlying processes we measured the saccadic eye movements observers made in this third experiment. Saccades serve the visual system by projecting interesting parts of the stimulus onto the fovea (the most sensitive central part of the retina). Between these saccades there are periods (called fixations) during which the eye does not move. Fixations usually last for 200 to about 300 ms (see Rayner, 1998). During fixations the intake of visual information may take place. The search time roughly depends on the number of fixations and the duration of these fixations. To get insight how search time is related to the number of fixations and fixation durations we first briefly discuss the processes underlying eye movement behaviour in search.

During fixations at least three processes are assumed to take place (Viviani, 1990). Firstly, the area around the fixation point is visually analysed (*foveal analysis*). This area from which visual (target) information is extracted is often referred to as the visual span (O'Regan, Lévy-Schoen, & Jacobs, 1983; Vlaskamp, Over, & Hooge, 2005). Secondly, the peripheral field is sampled for potential targets for the next saccade (*saccade target selection*). Thirdly, the next saccade is prepared; this takes about 150 ms (Becker & Jürgens, 1979).

The longer search times obtained in the displays containing 5% 2D non-targets may critically depend on two of these processes, namely *saccade target selection* and *foveal analysis*. If *saccade target selection* is hampered during search in the displays containing 5% 2D non-targets, we expect this to affect fixation locations (saccade landing positions). Note that saccade target selection is not the same as the selection of the target in a display. Saccade target selection is the selection of the potential targets to which the next saccade will be made. An indicator for hampered saccade target selection is the number of fixations on irrelevant red objects in relation to the number of irrelevant red objects in the display. As a result of hampered saccade target selection, we expect the number of fixations to be higher than in other conditions, caused by fixations on irrelevant red elements.

The second process, *foveal analysis* may also play a critical role in search behaviour. From the literature we know that when foveal analysis requires more time, it may lead to different eye movement strategies. For example, with decreasing target-distractor dissimilarity, the foveal analysis time increases (the visual system requires more time to detect the difference between a target and a distractor), which may lead to longer fixation times. Usually, fixation time increases with decreasing target-distractor dissimilarity (Hooge & Erkelens, 1996, 1998, 1999; Vlaskamp, Over, & Hooge, 2005). However, in dense displays (as in the present experiment) multiple stimulus elements are analysed in a single fixation. If foveal analysis requires more time this may affect behaviour in several ways. We describe two extremes: 1) Fixation time increases and is set in such way that the size of the area (or the number of elements) inspected remains constant, 2) Fixation time remains constant and the size of the inspected area (or the number of elements) decreases. These two strategies have different influences on the search time. Strategy 1 (fixation time increases, area inspected remains constant) produces longer search times through longer fixation times. Strategy 2 produces longer search times due to a higher number of fixations (the size of the area inspected decreases, a higher number of fixations are required to inspect the whole display) with an unchanged fixation time. Of course, a mix of these strategies may occur. Based on this mechanism, we expect fixation time and/or number of fixations to increase in the condition where displays contain 5% irrelevant non-targets, compared to the other conditions.

To examine the relation between saccades and search time, we repeated the first experiment, while measuring eye movements using the SR Eyelink II system.

Methods

Participants

Five observers (4 naïve as to the purpose of the experiment, 3 of them did not participate in experiment 1 or experiment 2) voluntarily participated in the experiment. All had normal or corrected to normal vision.

Stimuli, task and procedure

Stimuli, task and procedure for each subject was the same as in experiment 1. Stimuli were again presented on a Macintosh G4 computer using Matlab® 5.2.1, the Psychophysics Toolbox extensions (Brainard, 1997, Pelli, 1997) as well as the EyeLink Toolbox (Cornelissen, Peters & Palmer, 2002). However, in this experiment stimuli were presented on a 22 inch LaCie Electron Blue III monitor set to a resolution of 1280 by 1024 pixels and a vertical refresh rate of 75 Hz. Participants were placed in a chin rest such that the distance to the screen was 64 cm. This made stimulus size the same in terms of visual angle as in experiment 1. Each of the 17 2D non-target conditions (0-600) was repeated 20 times in each subject in two separate blocks containing a random order of 10 repetitions per condition. Each of the two experimental blocks started with a calibration (9 dots standard EyeLink calibration). Since we can record where observers fixated in the display, we can determine whether they found the target or not. Moreover, experiment 1 and 2 have shown that observers report the location of the target correctly.

Eye movement measurement

Movements of the left eye were measured at 500 Hz with the SR EyeLink II. Data were stored on disk and were analyzed off-line by self-written Matlab code.

Data analysis

The velocity signal of eye movements was searched for peak velocities above 20°/s. Each peak (in the velocity signal) was considered a potential indicator of the presence of a saccade. The exact onset of the saccade was determined by going backward in time to the point where the absolute velocity signal dropped below the average velocity plus two standard deviations during the stable fixation period before the saccade. The exact offset of the saccade was determined by going forward in time to the point where the absolute velocity signal dropped below the average velocity plus two standard deviations during the stable fixation period after the saccade (as in Vlaskamp, Over, & Hooge, 2005). This procedure was followed by rejection/acceptance based on minimum saccade duration of 10 ms and minimum amplitude of 1.0°. When a saccade was removed, fixation time before and after this saccade and the duration of the saccade were added together. Furthermore, based on search time outlier analysis, 4.23% of the data was discarded, since search time was 4.5 standard deviations (20 seconds or more, SD = 3.39 seconds) above the overall mean search time (5.06 seconds) in that part of the data. Such outliers were not present in experiment 1 and 2, and might be attributed to observers being distracted in some of the trials by the head-mounted Eyelink II in this third experiment.

Results and discussion

Search time and proportion fixated 2D non targets

Since this experiment was similar to experiment 1, in the present experiment, search time should show the same pattern as in experiment 1. Per observer the mean search time in each condition is calculated, after which the averages of these mean search times are calculated across subjects. The resulting mean search time is plotted in upper left panel of Figure 5.5. Immediately apparent is the increase in search time around 5% 2D non-targets, which is similar to that in the previous experiments (see grey area in the upper left panel of Figure 5.5).

As argued before, an increase in search time might be attributed to either a problem in *saccade target selection* or *foveal analysis* during search. An indicator for hampered selection is the number of fixations on irrelevant red objects in relation to the number of 2D irrelevant objects in the display. To compare across conditions (with different numbers of red elements), we computed proportions fixations on 2D non-targets and proportions of red elements in the display. This was done by dividing the number of fixations on 2D non-targets by the total number of fixations and by dividing the number of 2D non-targets by the total number elements in the display (600). Figure 5.6 shows proportion fixations on 2D non-targets versus proportions 2D non-targets in the display. Data points lie on the unity line, which indicates that the proportion

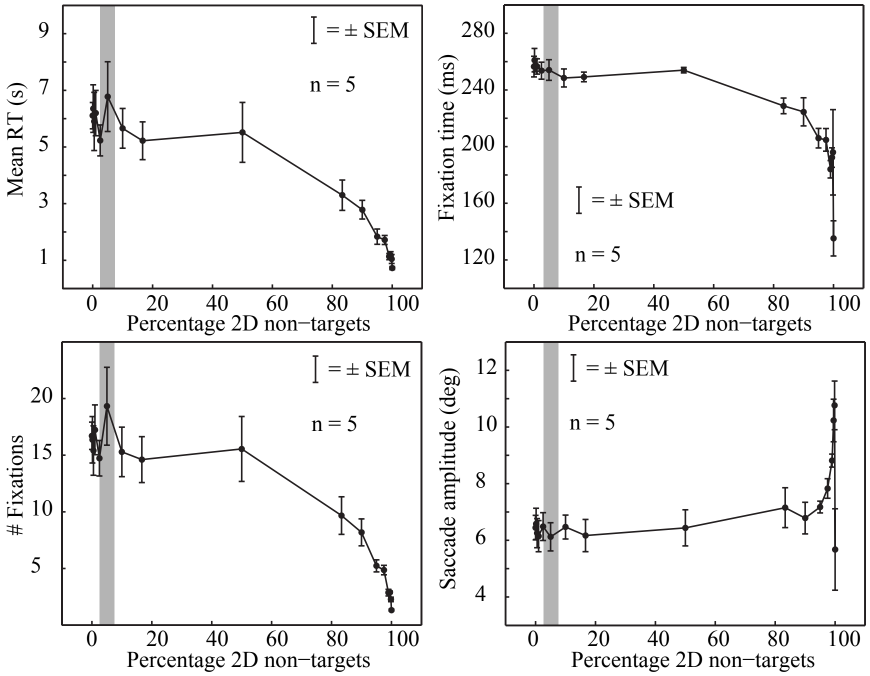

Figure 5.5. Average search time (RT in seconds) as a function of percentage 2D non-targets (upper left). Average fixation time (ms) as a function of percentage 2D non-targets (upper right). Average number of fixations as a function of percentage 2D non-targets (lower left). Average saccade amplitude (deg) as a function of percentage 2D non-targets (lower rightt). The gray areas indicate 2.5-7.5% 2D non-targets.

of fixations on red elements is equal to the proportion red elements in the display. In other words, the proportion fixations on 2D non-targets is directly related to the proportion 2D non-targets in the display. It is clear that saccade target selection is by no means affected by the size of the irrelevant 2D subset. From this we conclude that the longer search times obtained in displays containing 5% irrelevant items are not caused by hampered saccade target selection.

Fixation time, saccade amplitude and number of fixations

We have shown that hampered saccade target selection does not contribute to longer search time in the 5% 2D non-target condition, since 2D non-targets are not fixated more often than the proportion of 2D non-targets in the display would predict. As mentioned above, problems in the second process, *foveal analysis*, might contribute to longer search times in the 5% 2D non-target condition as well. We expect fixation time and/or number of fixations to increase (as expected based on the aforementioned strategies 1 and 2) in the condition where displays contain 5% irrelevant elements.

The average fixation time as a function of percentage 2D non-targets is constant up

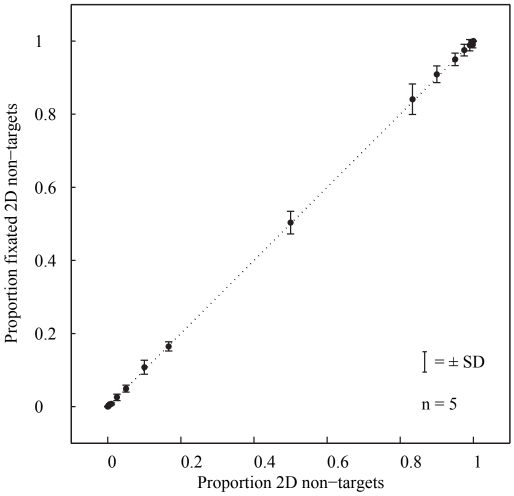

Figure 5.6. Average proportion fixated 2D non-targets as a function of proportion 2D non-targets present in the display. The dotted line indicates when proportion fixated 2D non-targets are equal to the proportion 2D non-targets present in the display. A proportion fixated 2D non-targets that falls above this line is in indication of not being able to avoid the 2D non-targets, while a proportion falling below this line it indicates that observers are able to select the relevant subset.

to a percentage of around 80% (see the upper right panel of Figure 5.5). The average fixation time at the 5% condition does not differ from average fixation time for a large range of conditions. When the percentage 2D non-targets is larger than 80%, fixation time decreases with increasing percentage of 2D non-targets. This pattern of fixation duration is not different for fixations on 2D or 1D non-targets ($F_{(1,30)} = 0.26$, $p = 0.60$). The shorter fixation time above 80% 2D non-targets might be attributed to the fact that the grey (relevant) 1D non-targets are spatially more separated across the display, making the smaller grey target easier to discriminate from grey 1D non-targets. This larger spacing between grey elements is reflected in saccadic amplitude as well. For a range of 2D non-target percentages up to 80% saccade amplitude (see the lower right panel of Figure 5.5) stays at the same level. Only when the 1D non-targets are spatially more separated across the display, the saccadic amplitude accordingly increases.

The other measure, namely the number of fixations, that could still explain the longer search time in the 5% 2D non-target condition is the number of fixations. As can be seen from the lower left panel of Figure 5.5 the number of fixations shows an almost identical pattern to that of search time. Apparently, a higher number of fixations are made in the 5% 2D non-target condition (again indicated by the grey area in lower left panel of Figure 5.5) compared to other conditions. This can also be seen in Figure 5.7, where the individual search time data is plotted against individual number of fixations data, individual fixation times and individual saccade amplitude. Search time correlates very strongly with the number of fixations ($r^2 = 0.984$ overall, $r^2 =$

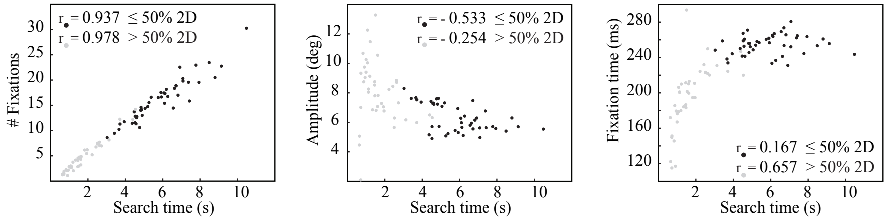

Figure 5.7. Individual search time data plotted against individual number of fixations (left panel), individual fixation durations (right panel) and individual saccade amplitudes (middle panel) and their correlations. Black dots indicate data taken from conditions where percentage 2D non-targets was equal or below 50%, gray dots indicate data from conditions where this percentage was above 50%.

0.937 for percentage 2D non-targets ≤ 50%, $r^2 = 0.978$, for percentage 2D > 50%), while search time correlates much weaker with fixation time ($r^2_{\text{overall}} = 0.754$, $r^2_{2D \leq 50\%} = 0.167$, and $r^2_{2D > 50\%} = 0.657$) and saccade amplitude ($r^2_{\text{overall}} = -0.544$, $r^2_{2D \leq 50\%} = -0.533$, and $r^2_{2D > 50\%} = -0.254$). Apparently, observers are searching longer when there is an irrelevant subset that is 5% of the total set size, since they make more fixations, while fixation time remains constant. This could be an indication of analysis of items taking more time in the 5% 2D non-target condition compared to other conditions.

To summarize, we have shown that longer search time in the 5% 2D non-target condition can not be attributed to hampered saccade target selection, since 2D non-targets are not fixated more often than the proportion of 2D non-targets in the display would predict. There does appear to be an overall longer foveal analysis in the 5% 2D non-target condition, which was indicated by a higher number of fixations in this condition, while fixation duration remained constant (the aforementioned strategy 2). Based on strategy 2 we assume that in the 5% 2D non-target condition apparently the visual span is decreased (a smaller area is inspected in the same fixation time). The question remains what causes this smaller visual span. It might be argued that covert mechanisms affect the visual processing of non-targets in the 5% non-target condition. This will be discussed in more detail in the General Discussion.

General Discussion

The current experiments set out to investigate whether observers were able to successfully select a relevant subset in a search display (with constant total set size), while the other irrelevant subset was varied in size. Alternatively, they might show an increase in search time when the irrelevant subset is small in relation to the total set size as was previously suggested by Wertheim (1981).

Based on the fact that this to be ignored irrelevant subset was clearly different in colour (equiluminant red vs. grey) observers should be able to select the grey relevant subset more easily (colour coded target, see Williams, 1966). The irrelevant non-targets differ on two dimensions (colour and size) from target, while the relevant non-targets differ on only one dimension (size). Therefore the first non-targets are referred to as 2D non-targets and the latter as 1D non-targets. Since total size of the search display was constant per trial, increasing the number of 2D non-targets (the irrelevant subset) means decreasing the size of the grey subset (the relevant subset), and as a consequence, search time for a grey target should decrease.

The large similarity between target and grey distractor (target was only 40% smaller in terms of degrees of visual angle) made the target hard to detect (cf. Duncan & Humphreys, 1989; Pashler, 1987). This large similarity should in turn yield a better selection of the grey subset (Hooge & Erkelens, 1999) due to longer fixations on possible targets. Again, when the grey subset is efficiently selected and more and more items are red, search time should decrease, since the grey subset becomes smaller as more 2D non-targets are added.

Interestingly, in all three experiments for a low percentage of 2D non-targets (around 5%) the search time is actually the longest compared to higher and lower percentages than that. In Experiment 1 the maximum search time for the grey target (in a display of 600 items) was not at zero percent 2D non-targets, but at 5%. Hence in experiment 2, which investigated whether this effect on search time was set size independent, four set sizes (96, 216, 360 and 600 items) were tested with only four levels of 2D non-targets (0, 5, 50 and 95%). Again, the 5% condition clearly showed longer search times compared to all other conditions as expressed by a larger search time slope for the 5% condition.

Experiment 3 showed results similar to experiment 1 and showed that search time patterns could not be attributed a *saccade target selection* problem as reflected in a proportion of 2D non-targets being fixated more often than was expected based on the proportion present in the display. Proportion fixated 2D non-targets followed proportion 2D non-targets in the display almost perfectly. This means that selecting the next potential target to make an eye movement to, while performing the search task, did not differ across percentages of irrelevant items.

Another possible explanation for longer search times at 5% 2D non-targets was that the time required for *foveal analysis* was longer in 5% 2D non-target conditions. This could both be reflected in fixation durations and/or in the number of fixations. However, fixation duration analysis showed no increase of fixation time around 5% irrelevant non-target conditions, nor did fixation time differ for 2D and 1D non-targets at any percentage of 2D non-targets up to 80% 2D non-targets. Note, that for percentages 2D non-targets above 80%, fixation time decreases and saccade amplitude increases. Only the number of fixations that observers made, which could be indicative of difficulties in analysing items currently fixated as well, did show a strong relationship with search time and showed a similar increase at 5% 2D non-targets.

The current results of all three experiments can be compared to results from Bacon and Egeth (1997), but only at percentages above 5% irrelevant non-targets. They show that search time results in their conjunction search task can be described by a quadratic bottom-up saliency-based distractor-ratio effect together with a linear decrease in search time due to top-down successful selection of a subset, decreasing in size. Fitting our data with this kind of function (a quadratic function centered around 50% 2D non-targets combined with a linear decrease as number of 2D non-targets increases) shows this can describe the data well at percentages at and above 10%. What remains striking though, is that at 5% 2D non-targets the data points can not be fitted well to a pattern as described by Bacon and Egeth (1997).

It can be argued that the extra increase in search time at 5% is the consequence of a bottom-up salience of the irrelevant subset. The smaller subset is more salient as argued by Bacon and Egeth (1997), but due to the fact that it is the red subset, the irrelevant subset becomes salient over the to be searched (grey) relevant subset and thus lengthens search time. In other words, in a non-conjunction search task as in the current experiment, a similar salience of the smaller subset results in a different search time pattern. Note however, that experiment 3 has shown that the irrelevant non-targets are not fixated more often than can be expected on the basis of the proportion of those non-targets in the display. Apparently, the irrelevant non-targets are not salient enough to yield overt orienting towards these items.

Alternatively, it is not the saliency per se of the irrelevant subset that causes longer search times. As was already proposed by Wertheim (1981), when the irrelevant red subset is much smaller than the grey subset, a red-grey figure-background percept might emerge. When only 5% is irrelevant, the target is not part of the figure, but is part of the background. Several studies have shown that is harder to find a target in the perceived background section of a stimulus pattern (Davis & Driver, 1998; He & Nakayama, 1992; Humphreys et al, 2000; Rensink & Enns, 1998). At 5% 2D non-targets this figure-ground segregation could be maximal and therefore increase search time maximally. Whether this figure-ground account really explains the current results and why this figure-ground percept is strongest at 5% irrelevant subset size should be tested in future research.

If fewer items per fixation are analysed, as the increase of number of fixations in experiment 3 suggests, it will take more time to find the target based on the probability of the target being present amongst the items currently fixated. This area in which items fall that are analysed during a fixation is sometimes referred to as visual span. Apparently visual span is decreased in the 5% 2D non-target condition. What causes this decrease in visual span at 5%, can, at this stage, only be speculated upon. A plausible option is that, for 5% irrelevant items, lateral masking of relevant items by the irrelevant items is maximal, since both a smaller or larger number of irrelevant items results in less

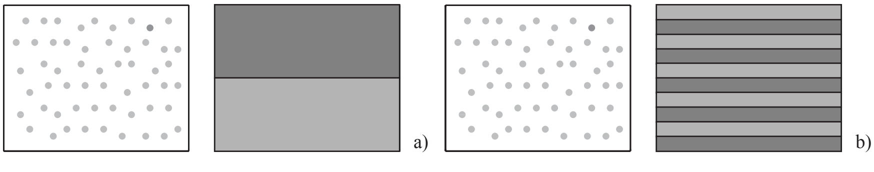

Figure 5.8. Example of search display (left), where a target has to be found (here indicated by the dark gray dot) with a) a low spatial distribution of ignore (light gray) and attend (dark gray) areas and b) a higher spatial distribution of these areas (cf. Gobell et al. 2004). The latter distribution of attention yields worse search performance than the first distribution.

crowding of relevant items. This could cause fewer items to be processed per fixation (Bouma, 1970; Wertheim, Hooge, Krikke, & Johnson 2006).

A recent article by Gobell and colleagues (2004) suggests that, when attending parts of a display, search performance depends on the spatial frequency with which attention is divided across the display. Dividing a search display into strips that should be ignored and ones that should be attended, the number of strips over which attention has to be divided determines how well search performance will be. Splitting up search displays in a small number of strips (a low spatial frequent distribution of attention, see Figure 5.8a) yields better search performance (more correct detections of targets when the display was presented 150 ms) than a display over which attention is distributed with a high spatial frequency (see Figure 5.8b). It could be that, up to 5% irrelevant items, irrelevant items might define strips/areas that can be ignored. As the percentage of irrelevant items increases to 5%, the strips/areas will become smaller (the spatial distribution of attention becomes higher in terms of spatial frequency). This will decrease search time. At percentages above 5%, the irrelevant items and relevant items might not be able to divide the display into attend and ignore areas any longer, but the irrelevant items might form a background on which the relevant items will stand out. The division of the display in attend and ignore areas might then become redundant and therefore search time decreases again when the percentage of irrelevant items increases above 5%.

In sum, we have shown that in feature search selection of the subset containing the target on the one hand and saliency of a to be ignored subset on the other hand interact in such a way that target detection is slowest with a small irrelevant subset, possibly by covert mechanisms decreasing visual span only when this irrelevant subset is small. Put differently, target detection critically depends on the number of irrelevant items.

Chapter 6

Summary and conclusions

As mentioned in the first chapter, this thesis set out to examine the limits of top-down attentional selection in perception. Although each sensory system has limits of its own (Kelly, 1979) and some of the sensory input will already be filtered, parts of this input needs to be selected (all visual input coming from all locations of the visual field, for example, can not be processed at the same time (Wolfe, 1994)). A selection that is based on current actions, goals, interests etcetera is needed to perform certain tasks adequately, while discarding the rest of task-irrelevant input. Such selections are made by means of attention and several models of how this selection takes place have been proposed and refined throughout the years (Rauschenberger & Yantis, 2006; Treisman & Gelade, 1980; Treisman & Gormican, 1988; Wolfe, Cave, & Franzel, 1989; Wolfe, 1994; Wolfe & Gancarz, 1996).

Besides describing *how* attention selects sensory input in such models, there is an extensive body of literature on *how well* attention can select sensory input. The spatiotemporal limits of this attentional selection are often described as if they were limits of different phenomena such as attentional blink (Shapiro, Raymond & Arnell, 1994), attentional capture (Jonides & Yantis, 1988; Hillstrom & Yantis, 1994), change blindness (Simons & Levin, 1997; Rensink, O'Regan & Clark, 2000), and banner blindness (Benway & Lane, 1998). It would be worthwhile, however, to describe these phenomena in the light of their common denominators, which are the temporal and spatial limits of attentional selection mechanisms and study those limits. In this thesis I set out to investigate how well attention top-down controls perception by examining these spatiotemporal limits of attention in perception.

In Chapter 2 I described how an apparent discrepancy between previous results concerned with the temporal limits of attentive tracking was solved (Horowitz, Holcombe, Wolfe, Arsenio & DiMase, 2004; Verstraten, Cavanagh & Labianca, 2000). Attentive tracking is based on the phenomenon of attention-based motion perception; attentionally selecting an object in a stimulus, which appears to move ambiguously (either rotating clockwise or counterclockwise) under passive viewing conditions, results in a percept of an unambiguously moving stimulus in one direction. This percept thus depends on the correct selection of objects in time and is therefore used to study temporal limits of attentional selection.

The apparent discrepancy was that one study (Verstraten et al. 2000) showed that objects needed to be separated in time for at least 70-100 ms to be tracked, while the other study (Horowitz et al. 2004) showed that objects needed to be separated 200 ms in time. I solved this discrepancy by showing that to be able to select the tracked object at its new location, observers need a signal to move their attention (a signal to disengage their attention from the present location and engage their attention to a new location). The signal that was needed for this disengagement was the offset (or disappearance) of the object. The timing of this offset signal turned out to be the critical determinant of being able to attentively track an object. By systematically varying what is called the duty cycle of the tracked object (the ratio between the interval when the object is present and absent in one location), I could vary the timing of this offset signal to disengage attention and reach a limit of shifting attention circa every 80 ms. The further the offset of the tracked object at the old location and the onset of that object at the new location were separated in time, the better tracking performance was. It was this offset signal in both studies that differed (co-occuring with next onset of the tracked object in Horowitz et al. (2004) and approximately 50 ms before next onset of the tracked object in Verstraten et al., (2000). This resulted in different temporal limits of attentional selection, both of which I could replicate by varying timing of the signal to disengage attention.

Another aspect that determines the temporal limits of attentional selection was described in Chapter 3. In this chapter I focussed on attentional selection of rapidly alternating simultaneously presented features. These features need to be combined into two alternating pairs of features. Being able to temporally select and integrate the correct features into pairs depends critically on the number of features over which integration has to take place in the visual domain (Holcombe and Cavanagh, 2001). As more features have to be selected and integrated into one percept, the separation in time between alternating features needs to be longer. I argued that sensory modality might be a stimulus feature for our brain as well, when selecting and integrating features into percepts. In Chapter 3, we found a decrease in temporal limits of cross-modal integration compared to unimodal visual and unimodal auditory integration in a series of experiments. I therefore argue that sensory modality is just another feature, which is an important factor in determining temporal limits of attentional selection.

Instead of investigating how well objects and events can be *selected*, we can also examine limits of attentional selection by investigating how well we can *ignore* irrelevant objects and events. We thus gain insight in how well we can attentionally select input in an indirect manner. If irrelevant input can not be ignored, then this should reflect the same limits of attentional selection that were investigated in Chapter 2 and 3. Therefore in Chapter 4 and 5 we examined attentional selection and its limits by studying what irrelevant events and objects can be ignored. In Chapter 4 I investigated temporal limits of attentional selection in this manner, while in Chapter 5 I focussed on spatial limits of attentional selection.

In Chapter 4 I investigated recent accounts (Lamy, 2005; Von Mühlenen, Rempel & Enns, 2005) on temporal properties of events that yield attentional capture. These two accounts seemed to contradict each other. Both explained why certain temporal properties of irrelevant events can capture our attention and thus disrupt the attentional selection of task-relevant objects and events. One of these accounts argued that any event that is temporally unique will capture attention (Von Mühlenen et al, 2005). Any event that does not co-occur with other events will yield attentional capture. The other account (Lamy, 2005) states that the timing of irrelevant events can be ignored, as long as the occurrence of these events is expected. Surprisingly, I found an overall different pattern of results that was consistent over two experiments. Based on my results I could refute the account of Von Mühlenen et al. (2005), as my results clearly demonstrate attentional capture by temporally non-unique events, However, closer inspection of

Von Mühlenen et al's raw reaction time data revealed that they showed even stronger instances of attentional capture by temporally non-unique events compared to results from our experiments. Furthermore, our results indicate that it was not the expected occurrence of irrelevant events (cf. Lamy, 2006), but the timing of the task-relevant target onset that determined whether observers could ignore an irrelevant event or not. Apparently, the specific timing of task-relevant events is yet another factor in determining the temporal limits of attentional selection.

In Chapter 5 I examined spatial limits of attentional selection in a series of visual search experiments. These experiments show that observers appear to be distracted by irrelevant non-targets in a large search display when the amount of these non-targets is small (5%). Search time is actually longer when there is small amount of irrelevant items (5% irrelevant items yields the overall longest search time) compared to any other amount of these items. In a second experiment we investigated whether this longest search time with 5% irrelevant items in a display was due to an absolute number of irrelevant items or due to a relative amount of irrelevant items. Independent of the total set size of the display, 5% irrelevant items yielded the longest search time compared to any other percentage of irrelevant items. Furthermore, the third experiment, using eye movement measurements, showed that the search time increasing effect of a small amount of irrelevant non-targets can not be attributed to a hampered saccade target selection when the irrelevant subset is 5% of the total set size. The proportion fixations on the irrelevant items was identical to the proportion of irrelevant items in the display and items were not fixated longer for this 5% condition. The only measure that correlated with search time appears to be the number of fixations. Since the total set size remained constant and all items (irrelevant and relevant) are fixated proportionally to their presence in a search display, it might be argued that these irrelevant non-targets decrease visual span maximally when they make up 5% of the non-targets. Visual span is defined as the area in which items are inspected during fixation. Apparently, attentional selection and visual span interact in these experiments. The precise mechanism by which this visual span is decreased by the small number of irrelevant non-targets remains unclear. An alternative explanation might be that above 5% irrelevant items, the irrelevant items might not be individuated by attention any longer (spatial distribution of attention can no longer be resolved) and the irrelevant items turn from individual noise items into a noise background or surface, which is easier to suppress. At this time, either of these explanations are based upon speculation. Further research is needed to determine which (if any) is the correct explanation.

New insights in temporal and spatial limits of attentional selection

As mentioned above, this thesis aimed at investigating *how well* we can top-down attentionally select objects and events from sensory input, both in time and across space. In the first chapter I briefly summarized what was known about both temporal and spatial limits of this attentional selection. After describing the experiments in chapter two to five, I can now compare outcomes of these results with these limits and further determine how well voluntarily control is exerted in perception.

If we consider what was known from research concerned with temporal limits of attentional selection, we were faced with a range of temporal resolutions of attention. Coming from different paradigms such as attentional blink (Chun & Potter, 1995; Raymond, Shapiro & Arnell, 1992; Duncan, Ward & Shapiro, 1994), attentional capture (Theeuwes, 2004; Theeuwes, De Vries & Godijn, 2003), attentive tracking (Horowitz, Holcombe, Wolfe, Arsenio & DiMase, 2004; Verstraten, Labianca, & Cavanagh, 2000), and synchrony judgments (Fujisaki & Nishida, 2005; Holcombe & Cavanagh, 2001), the sampling of sensory input by attention can be as fine as around sampling every 27 ms up to as coarse as sampling in intervals of 500 ms (though so called lag-1 sparing within the first 200 ms suggest a resampling of this 500 ms sampling in attentional blink paradigms). The experiments concerned with the temporal resolution of attention in this thesis seem to promote a rather fine temporal resolution of attention in Chapter 2 through 4. In Chapter 2 I have shown that attention can shift to a new location every 83 ms when tracking an object in a radially ambiguously apparent motion display. Attention can individuate events that are separated 83 ms in time. In Chapter 3 I have shown that to combine simultaneously alternating features into separately perceivable pairs, the feature pairs have to be separated 160 ms in time to be selected and integrated. Within one modality this separation between co-occuring features can be even smaller, shorter than circa 125 ms for auditory perception and shorter than 100 ms for visual perception. In Chapter 4 I show that there is no attentional capture when irrelevant events occur within 150 ms before target onset. Only when target and distractor events occur simultaneously with or 50 ms after target onset, attentional capture occurs. Apparently, attentional capture depends on the timing of the task-relevant event (target onset). This in contrast to previous results on attentional capture (Theeuwes, 2000) that show that attentional capture can recover when stimulus onset asynchrony (SOA) between target and irrelevant event is more than 150 ms. At shorter SOAs there should be attentional capture. This is referred to as the initial bottom-up sweep that can not be top-down controlled. In our experiments of Chapter 4, we have a range of SOAs that spans no more than 200 ms in our experiments. According to Theeuwes et al. (2000), we should find attentional capture at almost every SOA, which is clearly not the case. Thus it would be tempting to suggest that the initial bottom-up sweep is actually somewhat shorter. Taken together, in the experiments of Chapters 2-4 we find a temporal limit of attentional selection that lies around 80-100 ms, which is finer than coarse limits previously reported (150 to 300 ms).

An explanation for this discrepancy in the temporal limit of attentional selection might be attributed to the tasks under which these limits are measured. In attentional blink research observers need to report multiple task-relevant events/objects, while in attentional capture only one task-relevant event/object has to be reported and in attentive tracking one and the same relevant item has to be continuously monitored. It appears that the temporal limit of attention depends on what attention precisely needs to select. The simpler the required selection has to be, the finer the temporal resolution of attentional selection appears to be. Results from Chapter 3 appear to confirm this idea; in Chapter 2 to 4 we only need to select a certain number of task-relevant features.

When there are two task-relevant features to be selected, the temporal attentional selection limit is around 50-100 ms (tracking a *white object across multiple locations* in Chapter 2, matching *colour and contrast polarity* in unimodal visual conditions and *pitch to noise* in unimodal auditory conditions in Chapter 3, and finding the right *letter at a certain location* in Chapter 4) This temporal attentional selection limit becomes coarser (around 160 ms) as we have three task-relevant features (*matching colour, pitch, and sensory modality*) as in the cross-modal condition in Chapter 3.

Whether attention can individuate objects or events in time apparently depends on a couple of factors, which are the ability to disengage attention in time, the number of features that need to be selected simultaneously and the timing or occurrence of the task-relevant event/object itself.

For spatial limits of attentional selection I have shown in Chapter 5 that in search displays the time to find your target actually is longest when there is small number (5%) of irrelevant items. Search time is longest when 5% of the display is irrelevant, more or less irrelevant items both yield shorter search times.

Based on previous research on the spatial resolution of attention (Intriligator & Cavanagh, 2001), observers should be able to individuate items while searching the display. Observers are allowed to move their eyes across the display. This means that at every fixation items are right in the fovea, where spatial attentional resolution is about 3 arcminutes. Items spaced more than 3 arcminutes apart will be individuated by attention. However, this individuation decreases with eccentricity. At 15 degrees from the center of fixation, the critical spacing is around 2 degrees. Since, the items in our displays are separated 0.7 degrees on average, they should be individuated by attention at the center of fixation and a large area surrounding it. Following the equation by Intriligator and Cavanagh (2001) of the spatial threshold at a certain eccentricity with a doubling of the spatial limit every 0.38 degree items are separated further from the fovea, all items falling within an area of 10 degrees diameter should be individuated by attention with a spacing of 0.7 degrees.

A recent article by Gobell and colleagues (2004) might give an alternative explanation of the results from Chapter 5. They suggests that, when attending parts of a display, search performance depends on the spatial frequency with which attention is distributed across the display, rather than the spatial resolution of attention per se. It could be that, up to 5% irrelevant items, irrelevant items might determine strips/areas that can be ignored. As the percentage of irrelevant items increases to 5%, the strips/areas will become smaller (the spatial distribution of attention becomes higher in terms of spatial frequency). This will increase search time. At percentages above 5%, the irrelevant items and relevant items might not divide the display into attend and ignore areas any longer, but the irrelevant items might form a background on which the relevant items will stand out. The division of the display in *attend* and *ignore* areas might then become redundant and therefore search time decreases again when the percentage of irrelevant items increases above 5%.

Apparently, the spatial limits of attentional selection are influenced by the number of irrelevant items not purely by the spatial resolution of attention as defined by

Intriligator and Cavanagh (2001). Spatial attentional selection might depend on the spatial frequency with which attention has to be divided across the display.

In sum, I have shown in this thesis that the temporal limits of attentional selection are finer than thus far reported and that they depend on the ability to disengage attention, the number of features over which attention has to select, and the timing of the task-relevant events. For spatial limits of attentional selection the critical spacing between objects might not determine whether attention can select objects across space, but rather the spatial frequency with which attention is distributed across space determines spatial attentional selection in perception.

Nederlandse samenvatting

Stel je voor dat je in je auto zit. Het verkeerslicht staat op rood en je staat te wachten tot het groen wordt. Naast en achter je zitten je vrienden, druk in gesprek. Ze proberen jou te betrekken bij het gesprek, terwijl jezelf net werd afgeleid door een groen knipperende reclameverlichting aan de gevel van een winkelpand. Kortom, er gebeurt van alles om je heen, terwijl je taak is om je auto schadeloos van A naar B te krijgen. Met andere woorden, er is een continue stroom van binnenkomende zintuiglijke informatie, waaruit die informatie gehaald moet worden die relevant is voor de taak die je moet uit voeren.

Uit alles wat via onze zintuigen binnenkomt, moet dus een selectie gemaakt worden, terwijl de rest van die binnenkomende informatie genegeerd zal moeten worden. Hoewel het sensorische systeem zijn eigen beperkingen kent (Kelly, 1979) en sommige delen van deze binnenkomende sensorische informatie al gefilterd worden, is een verdere selectie van deze informatie noodzakelijk, aangezien zelfs ons brein een limiet kent aan de informatiehoeveelheid die het tegelijkertijd kan verwerken. Wolfe (1994) heeft bijvorbeeld aangetoond dat niet alle binnenkomende visuele informatie, die van alle locaties van het visuele veld komen, tegelijkertijd verwerkt kan worden. Een selectie van deze informatie is nodig op basis van huidige acties, doelen, interesses etcetera. Een dergelijke selectie wordt gemaakt door middel van aandacht.

Het doel van de experimenten in dit proefschrift is het onderzoeken van de beperkingen van deze aandachtsselectie in waarneming. Hoe deze aandachtselectie tot stand komt is in de loop der jaren in verscheidene modellen voorgesteld en meerdere malen aangescherpt (Rauschenberger & Yantis, 2006; Treisman & Gelade; 1980, Treisman & Gormican, 1988; Wolfe, Cave, & Franzel, 1989; Wolfe, 1994; Wolfe & Gancarz, 1996). Een belangrijke manier om te onderzoeken *hoe* door middel van aandacht deze selectie tot stand komt,is onderzoeken *hoe goed* binnenkomende sensorische informatie geselecteerd kan worden. Deze spatieel-temporele limieten van aandachtsselectie worden vaak beschreven alsof het limieten betreft van verschillende fenomenen zoals *attentional blink* (Shapiro, Raymond & Arnell, 1994), *attentional capture* (Jonides & Yantis, 1988; Hillstrom & Yantis, 1994), *change blindness* (Simons & Levin, 1997; Rensink, O'Regan & Clark, 2000), en *banner blindness* (Benway & Lane, 1998). Het is echter verstandiger deze fenomenen te beschrijven aan de hand van hun gemeenschappelijke kenmerk, namelijk de spatieel-temporele limieten van aandachtsselectie en deze limieten vervolgens te bestuderen. In dit proefschrift onderzoek ik daarom de mate van controle het aandachtsysteem over de waarneming uitoefent, door nu juist deze limieten te bestuderen.

In hoofdstuk 2 wordt beschreven hoe een schijnbare discrepantie kan worden opgelost tussen eerdere resultaten uit onderzoek, dat zich bezig houdt met de *temporele limieten van aandacht* aan de hand van *attentive tracking* (Horowitz, Holcombe, Wolfe, Arsenio & DiMase, 2004; Verstraten, Cavanagh & Labianca, 2000). Attentive tracking is gebaseerd op het fenomeen van aandachtsgebaseerde bewegingswaarneming: in een stimulus die onder passief bekijken ambigu lijkt te bewegen – namelijk in een richting met de klok mee of tegen de klok in – kan de waargenomen beweging veranderen in een schijnbare beweging die stabiel in één van de twee richtingen blijft bewegen. Dit is mogelijk door één object uit deze stimulus met aandacht te selecteren en te volgen. Deze aandachtgebaseerde waarneming hangt dus af van een correcte aandachtsselectie van objecten in de tijd en kan daarom gebruikt worden om de temporele limieten te bepalen van aandachtsselectie.

De schijnbare discrepantie was dat de ene studie (Verstraten et al. 2000) liet zien dat objecten, die met een interval van minimaal 70 tot 100 ms verschuiven, nog gevolgd kunnen worden met aandacht, terwijl de andere studie (Horowitz et al., 2004) stelde dat deze objecten slechts te volgen zijn wanneer zij met een interval van minimaal 200 ms verschuiven. Ik heb deze discrepantie kunnen oplossen door aan te tonen dat om een object op haar nieuwe locatie met aandacht te kunnen selecteren, aandacht een signaal nodig heeft om zich *los te maken* van de huidige locatie en te verschuiven naar de nieuwe locatie. Het signaal voor het los kunnen komen van de huidige locatie was het verdwijnen van het object op deze locatie. Het moment waarop dit verdwijnen zich voordoet, blijkt cruciaal in het op tijd kunnen verschuiven van aandacht naar de nieuwe locatie. Door het systematisch variëren van de levensduur van het gevolgde object (de verhouding tussen het aan- en afwezig zijn van een object op één en dezelfde locatie), kan het moment van verdwijnen van dit object gevarieerd worden, zodat de aandacht verschoven kan worden. Door dit te doen, kan een limiet bereikt worden, waarbij aandacht iedere 80 ms kan verschuiven naar een volgende locatie van een object. Hoe eerder het moment van verdwijnen van het object op de huidige locatie zich voordoet ten opzichte van het moment van verschijnen van het object op de nieuwe locatie, des te beter is het object te volgen. Dit moment van verdwijnen van het object, het zogenaamde *offset* signaal, was precies het signaal dat verschilde in beide bovengenoemde studies; het offset signaal deed zich gelijktijdig voor met het verschijnen van het gevolgde object op de nieuwe locatie (het *onset* signaal) in Horowitz et al. (2004), terwijl het offset signaal in Verstraten et al. (2000) 50 ms eerder plaats vond dan het onset signaal. Dit resulteerde in verschillende temporele limieten van aandachtselectie, die ik beide heb kunnen repliceren door het moment waarop aandacht zich los kan maken van een locatie te variëren.

Een ander aspect dat de temporele limieten van aandachtsselectie bepaalt, wordt beschreven in hoofdstuk 3. In dit hoofdstuk is gekeken naar de aandachtsselectie van elkaar snel afwisselende paren van stimuluseigenschappen die tot de waarneming van twee elkaar afwisselende objecten moet leiden. Om temporeel de juiste stimuluseigenschappen te kunnen selecteren en te integreren tot objecten hangt af van het aantal stimuluseigenschappen dat geïntegreerd moet worden. Dit is bekend voor integratie binnen het visuele domein (Bartels & Zeki, 2006; Holcombe & Cavanagh, 2001). Wanneer meer simultaan gepresenteerde stimuluseigenschappen gecombineerd moeten worden, dient het tijdsinterval tussen de verschillende paren van stimuluseigenschappen langer te zijn. Ik stelde dat sensorische modaliteit mogelijk ook een stimuluseigenschap is voor ons brein die meetelt in het selecteren en integreren van stimuli. Met sensorische modaliteit wordt de vorm (bijvoorbeeld auditief, visueel of tactiel) waarin de sensorische informatie binnenkomt bedoeld. Ik verwachtte dat de temporele limiet van integratie lager zou liggen (een grovere aandachtsselectie van stimuluseigenschappen in tijd) voor integratie *tussen* verschillende sensorische modaliteiten (cross-modaal) dan wanneer *binnen* één modaliteit (unimodaal) geïntegreerd moet worden. In hoofdstuk 3 werd inderdaad een afname in de temporele limiet gevonden wanneer cross-modale integratie werd vergeleken met unimodale visuele of unimodale auditieve integratie. Dit geeft mogelijk aan dat sensorische modaliteit inderdaad ook beschouwd kan worden als een stimuluseigenschap, die mede de temporele limiet van aandachtsselectie bepaalt.

In plaats van te onderzoeken hoe goed objecten en gebeurtenissen *geselecteerd* kunnen worden door aandacht, kunnen we aandachtsselectie ook onderzoeken door te bestuderen hoe goed irrelevante objecten en gebeurtenissen *genegeerd* kunnen worden. We verwerven dan hoe goed we met aandacht binnenkomende informatie kunnen selecteren op een indirecte manier. Als irrelevante informatie niet genegeerd kan worden, weerspiegelt dit mogelijk dezelfde limieten in aandachtsselectie, zoals bestudeerd in hoofdstuk 2 en 3. Daarom wordt aandachtselectie en de limieten daarvan onderzocht door te bestuderen welke irrelevante objecten en gebeurtenissen niet genegeerd kunnen worden. In hoofdstuk 4 onderzocht ik de temporele limieten van aandachtsselectie op deze indirecte manier, in hoofdstuk 5 lag de focus op spatiële limieten van aandachtsselectie.

In hoofdstuk 4 heb ik recente onderzoeken (Lamy, 2005; Von Mühlenen, Rempel & Enns, 2005) bestudeerd die zich bezig hielden met de temporele eigenschappen van irrelevante gebeurtenissen die leiden tot *attentional capture* (het 'vangen/trekken' of 'onderscheppen' van aandacht door taakirrelevante gebeurtenissen). Deze twee studies leken elkaar namelijk tegen te spreken. Beiden legden uit dat door bepaalde temporele eigenschappen van irrelevante gebeurtenissen aandachtselectie van taakrelevante gebeurtenissen en objecten verstoord kan worden. Eén van deze temporele verklaringen voor attentional capture was dat iedere gebeurtenis die *temporeel uniek* is de aandachtsselectie kan verstoren (Von Mühlenen et al., 2005); iedere gebeurtenis die zich niet gelijktijdig met de taakrelevante gebeurtenis voordoet, zal de aandacht trekken. De andere verklaring was dat het moment van het voordoen van een irrelevante gebeurtenis genegeerd kan worden, zolang deze irrelevante gebeurtenis zich op het *verwachte* moment voordoet. Ik vond een patroon van resultaten dat consistent naar voren kwam in een tweetal experimenten, maar dat anders was dan verwacht op basis van bovengenoemde temporele verklaringen. Op basis van mijn resultaten in hoofdstuk 4 kan de verklaring voor attentional capture door Von Mühlenen et al. (2005) verworpen worden, aangezien temporeel niet-unieke gebeurtenissen in de experimenten in hoofdstuk 4 toch leidden tot attentional capture. Een nadere inspectie van de ruwe reactietijddata van Von Mühlenen et al. (2005) laat echter ook zien dat attentional capture door temporeel niet-unieke gebeurtenissen zich ook voordoet in hun experimenten, zelfs met een sterker effect dan in mijn experimenten. Verder lijkt het niet de verwachting van het voordoen van een irrelevante gebeurtenis die het kunnen onderdrukken van attentional capture bepaalt (cf. Lamy, 2005), maar bepaalt het moment van voordoen van de taakrelevante gebeurtenis of irrelevante gebeurtenissen genegeerd kunnen worden of niet. Blijkbaar is het moment van voordoen van taakrelevante gebeurtenissen *nog* een factor die de temporele limieten van aandachtsselectie bepaalt.

In hoofdstuk 5 onderzocht ik spatiële limieten van aandachtsselectie in een serie experimenten op het gebied van visueel zoeken. De uitkomsten van deze experimenten lieten zien dat proefpersonen blijkbaar worden afgeleid door irrelevante elementen in een grote zoekplaat als de hoeveelheid van deze elementen klein was. De tijd om een doel te vinden, de zoektijd, was langer wanneer de hoeveelheid irrelevante elementen in een zoekplaat klein was (rond 5%) in vergelijking met situaties waar er meer of minder (of zelfs geen) irrelevante elementen in de zoekplaat zaten. In een tweede experiment is vervolgens aangetoond dat deze langste zoektijd met kleine hoeveelheid irrelevante elementen in een zoekplaat niet het gevolg was van een absoluut aantal irrelevante elementen, maar dat, onafhankelijk van de totale hoeveelheid elementen in een zoekplaat, de zoektijd het langst was als 5% van de elementen irrelevant is. Daarnaast hebben resultaten van een derde experiment, waarin oogbewegingen zijn gemeten, aangetoond dat het zoektijd verlengende effect van een kleine hoeveelheid irrelevante elementen niet toe te schrijven was aan een verhinderde saccadedoel-selectie (selectie van het element waar de volgende oogbeweging naartoe gemaakt zal worden), wanneer deze hoeveelheid 5% bedroeg. De proportie fixaties op irrelevante elementen was altijd gelijk aan de proportie irrelevante elementen in een zoekplaat. Verder werden elementen ook niet langer gefixeerd in deze 5% conditie. De enige maat die correleerde met de zoektijd bleek het aantal gemaakte fixaties te zijn. Aangezien de totale hoeveelheid elementen hetzelfde was in dit experiment en alle elementen (irrelevant en relevant) proportioneel ten opzichte van hun aanwezigheid in een zoekplaat gefixeerd werden, kan gesteld worden dat de irrelevante elementen de *visual span* verkleinen. Visual span is gedefinieerd als het gebied waarbinnen elementen geanalyseerd worden gedurende de duur van één fixatie. Blijkbaar interacteren aandachtsselectie en visual span in deze experimenten. Het precieze mechanisme achter deze verkleinde visual span bij een kleine hoeveelheid irrelevante elementen in een zoekplaat blijft onduidelijk. Een alternatieve verklaring is dat wanneer de hoeveelheid irrelevante elementen boven de 5% uitstijgt, de individuele irrelevante elementen niet meer door aandacht worden geselecteerd, waardoor de irrelevante elementen van individuele ruiselementen veranderen in één irrelevante ruisachtergrond of -oppervlak. Dit zou makkelijker te negeren zijn. Op dit moment zijn echter beide verklaringen slechts speculaties. Verder onderzoek is nodig om te bepalen welke van deze, of wellicht compleet andere, verklaringen juist is.

Nieuwe inzichten in de temporele en spatiële limieten van aandachtsselectie.

Zoals hierboven beschreven was het doel van deze dissertatie om te onderzoeken hoe goed met aandacht objecten en gebeurtenissen geselecteerd kunnen worden uit de binnenkomende sensorische informatie, zowel in tijd als in ruimte. In het inleidende eerste hoofdstuk is kort samengevat wat er al bekend is over de temporele en spatiële limieten van deze aandachtselectie. Na nu de experimenten beschreven te hebben uit hoofdstuk 2 tot en met 5, kan ik de uitkomsten van deze resultaten vergelijken met de bestaande kennis over temporele en spatiële limieten van aandachtsselectie en op die manier verder bepalen hoe goed uit binnenkomende sensorische informatie door aandacht geselecteerd kan worden.

Als we bekijken wat bekend is over de *temporele* limieten van aandachtsselectie, staan we voor het probleem dat de temporele resolutie van aandacht af lijkt te hangen van het gebruikte paradigma. *Attentional blink* (Chun & Potter, 1995; Raymond, Shapiro & Arnell, 1992; Duncan, Ward & Shapiro, 1994), *attentional capture* (Theeuwes, 2004; Theeuwes, De Vries & Godijn, 2003), *attentive tracking* (Horowitz, Holcombe, Wolfe, Arsenio & DiMase, 2004; Verstraten, Labianca, & Cavanagh, 2000), en *synchrony judgments* (Fujisaki & Nishida, 2005; Holcombe & Cavanagh, 2001) paradigma's suggereren dat het peilen van sensorische informatie door aandacht zo precies als iedere 27 ms tot zo grof als iedere 500 ms plaats vindt. De experimenten uit deze dissertatie die zich bezig houden met de temporele resolutie van aandacht lijken te wijzen op een redelijk precieze temporele resolutie in hoofdstuk 2 tot en met 4. In hoofdstuk 2 heb ik laten zien dat aandacht iedere 83 ms naar een nieuwe locatie verschoven kan worden, wanneer een object gevolgd wordt in een anderszins ambigu bewegende stimulus. Aandacht kan in dit geval individuele gebeurtenissen onderscheiden als deze 83 ms van elkaar gescheiden zijn. In hoofdstuk 3 laat ik zien dat om simultaan gepresenteerde paren van elkaar afwisselende stimuluseigenschappen waar te nemen als twee elkaar afwisselende gehele stimuli, deze simultaan gepresenteerde paren van stimuluseigenschappen 160 ms van elkaar gescheiden moeten zijn. Wanneer deze selectie binnen één modaliteit plaatsvindt, kan de scheiding tussen paren stimuluseigenschappen zelfs nog korter zijn. Voor auditieve waarneming kan deze scheiding in tijd 125 ms zijn en voor visuele waarneming kan deze scheiding zelfs kleiner dan 100 ms zijn. In hoofdstuk 4 laat ik zien dat er zich geen attentional capture voordoet, wanneer irrelevante gebeurtenissen zich voordoen in de 150 ms die vooraf gaan aan het verschijnen van een te rapporteren doel. Slechts wanneer dit doel verschenen is, net als 50 ms na het verschijnen van het doel, veroorzaken irrelevante gebeurtenissen attentional capture. Dit spreekt resultaten uit eerder gedaan onderzoek (Theeuwes et al., 2000) tegen, waaruit blijkt dat effecten van attentional capture teniet gedaan kunnen worden als de tijd tussen verschijnen van het doel en het optreden van de irrelevante gebeurtenis (de zogenaamde stimulus onset asynchrony, afgekort SOA) langer is dan 150 ms. Bij kortere SOA's zou attentional capture zich moeten voordoen. Dit tijdsinterval wordt vaak beschreven als een initiële *bottom-up sweep* van een irrelevante gebeurtenis die niet door 'hogere' processen gecontroleerd kan worden. In de experimenten van hoofdstuk 4 is een SOA-bereik gebruikt dat in zijn geheel niet langer is dan 200 ms. Volgens Theeuwes et al. (2000) zou zich daarom bij bijna iedere SOA die gebruikt is in de experimenten van hoofdstuk 4 attentional capture moeten voordoen. Dit is echter niet het geval. Het zou daarom zo kunnen zijn dat de initiële bottom-up sweep korter is dan tot dusver gerapporteerd. Als we de resultaten van de experimenten uit hoofdstuk 2 tot en met 4 naast elkaar leggen, kan gesteld worden dat de temporele limiet van aandachtsselectie zo rond de 80 tot 100 ms ligt, wat preciezer is dan gerapporteerd in eerder onderzoek (150 tot 500 ms).

Een verklaring voor deze discrepantie in temporele limieten van aandachtsselectie zou kunnen liggen in de specifieke taken waarin deze limieten gemeten zijn. In *attentional blink* onderzoek dienen proefpersonen meerdere taakrelevante gebeurtenissen/objecten te rapporteren, terwijl in *attentional capture* paradigma's slechts één taakrelevante gebeurtenis/object gerapporteerd hoeft te worden. In *attentive tracking* paradigma's kan volstaan worden met het continu monitoren van een object. Blijkbaar hangt de temporele limiet van aandachtsselectie af van wat precies geselecteerd moet worden. Hoe simpeler de benodigde selectie kan zijn, hoe preciezer de temporele resolutie van aandachtsselectie blijkt te zijn. De resultaten uit hoofdstuk 3 lijken dit idee te bevestigen: in hoofdstuk 2 tot en met 4 moeten steeds slechts telkens een aantal taakrelevante stimuluseigenschappen geselecteerd worden. Wanneer er *twee* taakrelevante stimuluseigenschappen geselecteerd moeten worden, ligt de temporele limiet van aandachtsselectie rond de 50 tot 100 ms (het volgen van een *wit object* langs verschillende *locaties* in hoofdstuk 2, het koppelen van *kleur* en *contrast polariteit* in unimodale visuele condities en *toonhoogte* en *ruis* in unimodale auditieve condities in hoofdstuk 3, en het vinden van de juiste *letter* op een bepaalde *locatie* in hoofdstuk 4). Deze temporele limiet van aandachtsselectie wordt minder nauwkeurig (rond 160 ms) als er *drie* taakrelevante stimuluseigenschappen geselecteerd dienen te worden (*kleur*, *toonhoogte* en *sensorische modaliteit*) in de cross-modale conditie in hoofdstuk 3.

Of aandacht individuele objecten en gebeurtenissen nog in tijd kan scheiden en selecteren blijkt dus af te hangen van een aantal factoren: de mogelijkheid om aandacht op tijd los te maken om deze vervolgens te kunnen verschuiven, het aantal stimuluseigenschappen dat tegelijkertijd geselecteerd dient te worden en het moment van voordoen van de taakrelevante gebeurtenis/object zelf.

Wat de *spatiële limiet* van aandachtsselectie betreft heb ik in hoofdstuk 5 laten zien dat de tijd om een doel te vinden in een zoekplaat het langste is wanneer er een kleine hoeveelheid irrelevante elementen in de zoekplaat aanwezig is. Zoektijd is het langste wanneer de zoekplaat 5% irrelevante elementen bevat. Meer of minder irrelevante elementen leveren beide kortere zoektijden op.

Gebaseerd op eerder onderzoek dat zich bezig heeft gehouden met de spatiële resolutie van aandacht (Intriligator & Cavanagh, 2001), zouden proefpersonen in staat moeten zijn om de individuele elementen te kunnen onderscheiden in de zoekplaten. Het staat de proefpersonen vrij oogbewegingen te maken, terwijl de zoekplaat doorzocht wordt. Dit betekent dat bij iedere fixatie elementen zich direct in de fovea bevinden, waar de spatiële resolutie van aandacht rond de 3 boogminuten ligt. Elementen die meer dan 3 boogminuten van elkaar gescheiden zijn zullen individueel geselecteerd worden aandacht. Echter de precisie in het kunnen onderscheiden door aandacht van individuele elementen neemt af met eccentriciteit ten opzichte van het centrum van fixatie. Op 15 graden vanuit het centrum van fixatie is de kritieke ruimte, die tussen elementen moet zitten zodat ze met aandacht nog te onderscheiden zijn, 2 graden. In de zoekplaten van hoofdstuk 5 zijn de elementen gemiddeld 0.7 graden van elkaar gescheiden. Aan de hand van de berekening voor de kritieke ruimte tussen elementen die Intriligator en Cavanagh (2001) voorstellen, waarbij de spatiële limiet bij iedere 0.38 graden vanuit het centrum van fixatie verdubbelt, zouden individuele elementen die binnen een gebied met een diameter van 10 graden vallen door aandacht geselecteerd kunnen worden.

Een recent artikel van Gobell en collega's (2004) geeft een mogelijke andere verklaring voor de resultaten van hoofdstuk 5. Zij stellen dat, wanneer aandacht verdeeld is over verschillende delen van een zoekplaat, de zoekprestatie afhangt van de spatiële frequentie waarmee aandacht verdeeld is over de zoekplaat, in tegenstelling tot de spatiële resolutie van aandacht op zich. Het zou kunnen zijn dat – tot aan situaties met 5% irrelevante elementen – de irrelevante elementen gebieden definiëren die genegeerd kunnen worden. Hoe groter het percentage irrelevante elementen wordt in een zoekplaat, des te kleiner worden de te negeren gebieden. Dit is hetzelfde als een spatiële verdeling van aandacht met een hoger spatiële frequentie. Dit zal de zoektijd verlengen. Echter bij percentages boven de 5% zou het kunnen zijn dat irrelevante en relevante elementen geen gebieden meer definiëren die genegeerd dan wel geattendeerd moeten worden. De irrelevante elementen zouden dan mogelijk een achtergrond kunnen gaan vormen waartegen de relevante elementen zich aftekenen. De verdeling van de zoekplaat in negeer- en attendeergebieden wordt op dat moment overbodig. Daarom zou de zoektijd weer kunnen afnemen wanneer een zoekplaat meer dan 5% irrelevante elementen bevat.

Blijkbaar hangt de spatiële limiet van aandachtsselectie af van de hoeveelheid irrelevante elementen, maar niet op basis van de spatiële resolutie van aandacht zoals gedefinieerd door Intriligator en Cavanagh (2001). Spatiële aandachtsselectie hangt mogelijk af van de spatiële frequentie van verdeling van aandacht in een zoekplaat.

Bij elkaar genomen heb ik in deze dissertatie laten zien dat de temporele beperkingen in aandachtsselectie kleiner zijn dan tot dusver waren gerapporteerd. Deze beperkingen hangen af van de mogelijkheid om aandacht los te maken, van het aantal stimuluseigenschappen dat door aandacht geselecteerd dient te worden en het moment van voordoen van taakrelevante gebeurtenissen. Wat betreft de spatiële beperkingen van aandachtsselectie bepaalt niet de kritieke ruimte tussen elementen de spatiële aandachtselectie, maar bepaalt eerder de spatiële frequentie waarmee aandacht is verdeeld over de ruimte deze spatiële aandachtsselectie.

References

Alais, D., & Carlile, S. (2005). Synchronizing to real events: subjective audiovisual alignment scales with perceived auditory depth and speed of sound. Proceedings of the National Academy of Sciences, 102, 2244-2247.

Arnold, D.H., Johnston, A., & Nishida, S. (2005). Timing sight and sound. Vision Research, 45, 1275-1284.

Arrighi, R., Alais, D., & Burr, D. (2006). Perceptual synchrony of audiovisual streams for natural and artificial motion sequences. Journal of Vision, 6, 260-268.

Bacon, W.J., & Egeth, H.E. (1997). Goal-directed guidance of attention: evidence from conjunctive visual search. Journal of Experimental Psychology: Human Perception and Performance, 23, 948-961.

Bartels, A., & Zeki, S. (2006). The temporal order of binding visual attributes. Vision Research, 46, 2280-2286.

Becker, W., & Jürgens, R. (1979). An analysis of the saccadic system by means of double step stimuli. Vision Research, 19, 967-83.

Benway, J.P., & Lane, D.M. (1998). Banner blindness: web searchers often miss the "obvious" links. Internetworking, 1.3

Bodelón, C., Fallah, M., & Reynolds, J.H. (2007). Temporal resolution for the perception of features and conjunctions. Journal of Neuroscience, 27, 725-730.

Boer, E. de (1985). Auditory time constants: a paradox? In: A Michelsen (Ed.), Time resolutions in auditory systems, Berlin: Springer-Verlag.

Bouma, H. (1970). Interaction effects in parafoveal letter recognition. Nature, 226, 177-178.

Bouma, H. (1973). Visual interference in the parafoveal recognition of initial and final letters of words. Vision Research, 13, 767-782.

Brainard, D.H. (1997). The Psychophysics Toolbox. Spatial Vision, 10, 433-436.

Carrasco, M., Giordano, A.M., & McElree, B. (2004). Temporal performance fields: visual and attentional factors. Vision Research, 44, 1351-1365.

Carrasco, M., Ling, S., & Read, S. (2004). Attention alters appearance. Nature Neuroscience, 7, 308-313.

Cavanagh, P. (1992). Attention-based motion perception. Science, 257, 1563-1565.

Cavanagh, P., Tyler, C.W., & Favreau, O.E. (1984). Perceived velocity of moving chromatic gratings. Journal of the Optical Society of America. A, optics and image science, 1, 893-899.

Chun, M.M., & Potter, M.C. (1995). A two-stage model for multiple target detection in rapid serial visual presentation. Journal of Experimental Psychology: Human perception and performance, 21, 109-127.

Cornelissen, F.W., Peters, E.M., & Palmer, J. (2002). The Eyelink Toolbox: Eye tracking with MATLAB and the Psychophysics Toolbox. Behavior, Research Methods, Instruments, & Computers, 34, 613-617.

Danckert, J., & Maruff, P. (1997). Manipulating the disengage operation of covert visual spatial attention. Perception and Psychophysics, 59, 500-508.

Davis, G., & Driver, J. (1998). Kanisza subjective figures can act as occluding surfaces at parallel stages of visual search. Journal of Experimental Psychology: Human Perception and Performance, 24,169-84.

Dixon, N.F., & Spitz, L. (1980). The detection of auditory visual desynchrony. Perception, 9, 719-721.

Duncan, J., & Humphreys, G.W. (1989). Visual search and stimulus similarity. Psychological Review, 96, 433-458.

Duncan, J., Ward, R., & Shapiro, K. (1994). Direct measurement of attentional dwell time in human vision. Nature, 369, 313-315.

Egeth, H., & Dagenbach, D. (1991). Parallel versus serial processing in visual search: further evidence from subadditive effects of visual quality. Journal of Experimental Psychology: Human Perception and Performance, 17, 551-560.

Eriksen, C.W., & Hoffman, J.E. (1973). The extent of processing of noise elements during selective encoding from visual displays. Perception and Psychophysics, 14, 155-160.

Eriksen, C.W., & St James, J.D. (1986). Visual attention within and around the field of focal attention: a zoom lens model. Perception and Psychophysics, 40, 225-240.

Fendrich, R., & Corballis, P.M. (2001). The temporal cross-capture of audition and vision. Perception and Psychophysics, 63, 719-725.

Fischer, B., & Ramsperger, E. (1984). Human express saccades: extremely short reaction times of goal directed eye movements. Experimental Brain Research, 57, 191-195.

Folk, C.L., & Remington, R. (1998). Selectivity in distraction by irrelevant featural singletons: evidence for two forms of attentional capture. Journal of Experimental Psychology: Human Perception and Performance, 24, 847-858.

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., & Wright, J.H. (1994). The structure of attentional control: contingent attentional capture by apparent motion, abrupt onset, and color. Journal of Experimental Psychology: Human Perception and Performance, 24, 317-329.

Foster, D.H., & Ward, P.A. (1991). Asymmetries in oriented-line detection indicate two orthogonal filters in early vision. Proceedings of the Royal Society, Series B: Biological Sciences, 243, 75-81.

Fujisaki, W., & Nishida, S. (2005). Temporal frequency characteristics of synchrony-asynchrony discrimination of audiovisual signals. Experimental Brain Research, 166, 455-464.

Fujisaki, W., Koene, A., Arnold, D.H., Johnston. A., & Nishida, S. (2006). Visual search for a target changing in synchrony with an auditory signal. Proceedings of the Royal Society London Series B: Biological Sciences, 273, 865-874.

Fujisaki, W., Shimojo, S., Kashino, M., & Nishida, S. (2004). Recalibration of audio-visual simultaneity. Nature Neuroscience, 7, 773-778.

Gebhard, J.W., Mowbray, G.H. (1959). On discriminating the rate of visual flicker and auditory flutter. American Journal of Psychology, 72, 521-529.

Gobell, J.L., Tseng, C.H., & Sperling, G. (2004). The spatial distribution of visual attention. Vision Research, 44, 1273-1296.

Green, D.M. (1985). Temporal factors in psychoacoustics. In A Michelsen (Ed.), Time resolutions in auditory systems, Berlin: Springer-Verlag.

He, Z., & Nakayama, K. (1992). Surfaces versus features in visual search. Nature, 359, 231-233.

Hillstrom, A.P., & Yantis, S. (1994). Visual motion and attentional capture. Perception and Psychophysics, 55, 399-411.

Holcombe, A.O., & Cavanagh, P. (2001). Early binding of feature pairs for visual perception. Nature Neuroscience, 4, 127-128

Hooge, I.T., & Erkelens, C.J. (1996). Control of fixation duration in a simple search task. Perception and Psychophysics, 58, 969-976.

Hooge, I.T., & Erkelens, C.J. (1998). Adjustment of fixation duration in visual search. Vision Research, 38, 1295-1302.

Hooge, I.T., & Erkelens, C.J. (1999). Peripheral vision and oculomotor control during visual search. Vision Research, 39, 1567-1575.

Horowitz, T.S., Holcombe, A.O., Wolfe, J.M., Arsenio, H.C., & DiMase, J.S. (2004). Attentional pursuit is faster than attentional saccade. Journal of Vision, 4, 585-603.

Humphreys, G.W., & Müller, H. (2000). A search asymmetry reversed by figure-ground assignment. Psychological Science, 11, 196-201.

Intriligator, J., & Cavanagh, P. (2001). The spatial resolution of visual attention. Cognitive Psychology, 43, 171-216.

James, W. (1890). Principles of psychology (vol. 1). New York: Holt (Reprinted in 1950 by Dover Press, New York).

Johnston, A., & Nishida, S. (2001). Time perception: brain time or event time? Current Biology, 11, R427-430.

Jonides, J., & Yantis, S. (1988). Uniqueness of abrupt visual onset in capturing attention. Perception and Psychophysics, 43, 346-354.

Kahneman, D., & Treisman, A. (1984). Changing views of attention and automaticity. In: D.R.D. Parasuraman (Ed.), Varieties of attention, Orlando: Academic Press.

Kahneman, D., Treisman, A., & Gibbs, B. J. (1992). The reviewing of object files: Object-specific integration of information. Cognitive Psychology, 24, 175-219.

Keetels, M., & Vroomen, J.,(2005). The role of spatial disparity and hemifields in audio-visual temporal order judgments. Experimental Brain Research, 167, 635-640.

Kelly, D.H. (1979). Motion and vision. II. Stabilized spatio-temporal threshold surface. Journal of the Optical Society of America, 69, 1340-1349.

Klein, R.M., Taylor, T.L., & Kingstone, A. (1995). Against a role for attentional disengagement in the gap effect: a friendly amendment to Tam and Stelmach (1993). Perception and Psychophysics, 57, 573-577.

Lamy, D. (2005). Temporal expectations modulate attentional capture. Psychonomic Bulletin and Review, 12, 1112-1119.

Lamy, D., & Egeth, H.E. (2003). Attentional capture in singleton-detection and feature-search modes. Journal of Experimental Psychology: Human Perception and Performance, 29, 1003-1020.

Lamy, D., Leber, A., & Egeth, H.E. (2004). Effects of task relevance and stimulus-driven salience in feature-search mode. Journal of Experimental Psychology: Human Perception and Performance, 30, 1019-1031.

Leber, A.B., & Egeth, H.E. (2006). It's under control: top-down search strategies can override attentional capture. Psychonomic Bulletin and Review, 13, 132-138.

Lewald, J., & Guski, R. (2003). Cross-modal perceptual integration of spatially and temporally disparate auditory and visual stimuli. Cognitive Brain Research, 16, 468-478.

Lewkowicz, D.J. (1996). Perception of auditory-visual temporal synchrony in human infants. Journal of Experimental Psychology: Human Perception and Performance, 22 (5), 1094-1106.

Luria, S.M., & Strauss, M.S. (1975). Eye movements during search for coded and uncoded targets. Perception and Psychophysics, 17, 303-308.

Mackeben, M., & Nakayama, K. (1993). Express attentional shifts. Vision Research, 33, 85-90.

Maljkovic, V. & Nakayama, K. (1994). Priming of popout: I Role of features. Memory and Cognition, 22, 657-672.

Morein-Zamir, S., Soto-Faraco, S., & Kingstone, A. (2003). Auditory capture of vision: examining temporal ventriloquism. Cognitive Brain Research, 17, 154-163.

Moutoussis, K., & Zeki, S. (1997). A direct demonstration of perceptual asynchrony in vision. Proceedings of the Royal Society of London Series B: Biological Sciences, 264, 393-399.

Mühlenen, A. von, Rempel, M.I., & Enns, J.T. (2005). Unique temporal change is the key to attentional capture. Psychological Science, 16, 979-986.

Neisser, U. (1963). Decision time without reaction time: Experiments in visual scanning. American Journal of Psychology, 76, 376-385

Nishida, S., & Johnston, A. (2002). Marker correspondence, not processing latency, determines temporal binding of visual attributes. Current Biology, 12, 359-368.

Noesselt, T., Fendrich, R., Bonath, B., Tyll, S., & Heinze, H.J. (2005). Closer in time when farther in space - spatial factors in audiovisual temporal integration. Cognitive Brain Research, 25, 443-458.

O'Regan, J.K., Lévy-Schoen, A., & Jacobs, A.M. (1983). The effect of visibility on eye-movement parameters in reading. Perception and Psychophysics, 34, 457-464.

Pashler, H. (1987). Target-distractor discriminability in visual search. Perception and Psychophysics, 41, 385-392.

Pelli, D.G. (1997). The VideoToolbox software for visual psychophysics: transforming numbers into movies. Spatial Vision, 10, 437-442.

Posner, M.I. (1978). Chronometric explorations of the mind. Hillsdale, NJ: Erlbaum.

Posner, M.I. (1980). Orienting of attention. Quarterly Journal of Experimental Psychology, 32, 3-25.

Posner, M.I., Snyder, C.R., & Davidson, B.J. (1980). Attention and the detection of signals. Journal of Experimental Psychology, 109, 160-174.

Pylyshyn, Z. (1989). The role of location indexes in spatial perception: a sketch of the FINST spatial-index model. Cognition, 32, 65-97.

Rauschenberger, R. (2003). Attentional capture by auto- and allo-cues. Psychonomic Bulletin and Review, 10, 814-842.

Rauschenberger, R., & Yantis, S. (2006). Perceptual encoding efficiency in visual search. Journal of Experimental Psychology: General, 135, 116-131.

Raymond, J.E., Shapiro, K.L., & Arnell, K.M. (1992). Temporary suppression of visual processing in an RSVP task: an attentional blink? Journal of Experimental Psychology: Human Perception and Performance, 18, 849-860.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. Psychological Bulletin, 124, 372-422.

Recanzone, G.H. (2003). Auditory influences on visual temporal rate perception. Journal of Neurophysiology, 89, 1078-1093.

Rensink, R., & Enns J.T. (1998). Early completion of occluded objects. Vision Research, 38, 2489-2505.

Rensink, R., O'Regan, J., & Clark, J. (2000). On the failure to detect changes in scenes across brief interruptions. Visual Cognition, 7, 127-145.

Ruz, M., & Lupianez, J. (2002). A review of attentional capture: on its automaticity and sensitivity to endogenous control. Psicológica, 23, 283-309.

Sagi, D., & Julesz, B. (1986). Enhanced detection in the aperture of focal attention during simple discrimination tasks. Nature, 321, 693-695.

Shapiro, K.L., Raymond, J.E., & Arnell, K.M. (1994). Attention to visual pattern information produces the attentional blink in rapid serial visual presentation. Journal of Experimental Psychology: Human Perception and Performance, 20, 357-371.

Shen, J., Reingold, E.M., & Pomplun, M. (2003). Guidance of eye movements during conjunctive visual search: the distractor-ratio effect. Canadian Journal of Experimental Psychology, 57, 76-96.

Shipley T, (1964). Auditory flutter-driving of visual flicker. Science, 145, 1328-1330.

Shore, D.I., Spence, C., & Klein, R.M. (2001). Visual prior entry. Psychological Science, 12, 205-212.

Simons, D.J., & Levin, D.T. (1997). Change blindness. Trends in Cognitive Sciences, 1, 261-267.

Sugita, Y., & Suzuki, Y. (2003). Implicit estimation of sound-arrival time. Nature, 421, 911.

Tam, W.J., & Stelmach, L.B. (1993). Viewing behavior: ocular and attentional disengagement. Perception and Psychophysics, 54, 211-222.

Theeuwes, J. (1990). Perceptual selectivity is task dependent: evidence from selective search. Acta Psychologica, 74, 81-99.

Theeuwes, J. (1995). Abrupt luminance change pops-out; Abrupt color change does not. Perception and Psychophysics, 57, 637-644.

Theeuwes, J., Atchley, P. & Kramer, A.F. (2000). On the time course of top-down and bottom-up control of visual attention. In: S. Monsell, & J. Driver (Eds.), Attention &

Performance (Vol 18). Cambridge: MIT Press.

Theeuwes, J., De Vries, G.J., & Godijn, R. (2003). Attentional and oculomotor capture with static singletons. Perception and Psychophysics, 65, 735-746.

Theeuwes, J. (2004). Top-down search strategies cannot override attentional capture. Psychonomic Bulletin and Review, 11, 65-70.

Theeuwes, J. & Chen, C.Y.D (2005). Attentional capture and inhibition (of return): the effect on perceptual sensitivity. Perception and Psychophysics, 67, 1305-1312.

Toet, A., & Levi, D.M. (1992). The two-dimensional shape of spatial interaction zones in the parafovea. Vision Research, 32, 1349-1357.

Treisman, A., & Gormican, S. (1988). Feature analysis in early vision: evidence from search asymmetries. Psychological Review, 95, 15-48.

Treisman, A., & Sato, S. (1990). Conjunction search revisited. Journal of Experimental Psychology: Human Perception and Performance, 16, 459-478.

Treisman, A.M., & Gelade, G. (1980). A feature-integration theory of attention. Cognitive Psychology, 12, 97-136.

Tse, P.U., Sheinberg, D.L., & Logothetis, N.K. (2003). Attentional enhancement opposite a peripheral flash revealed using change blindness. Psychological Science, 14, 91-99.

Verstraten, F.A., Cavanagh, P., & Labianca, A.T. (2000). Limits of attentive tracking reveal temporal properties of attention. Vision Research, 40, 3651-3664.

Viemeister, N.F., & Wakefield, G.H. (1991). Temporal integration and multiple looks. Journal of the Acoustic Society of America, 90, 858-865.

Viviani, P. (1990). Eye movements in visual search: cognitive, perceptual and motor control aspects. In: E. Kowler (Ed.), Reviews of oculomotor research: eye movements and their role in visual and cognitive processes. Amsterdam: Elsevier.

Vlaskamp, B.V., Over, E.A., & Hooge, I.T. (2005). Saccadic search performance: the effect of element spacing. Experimental Brain Research, 167, 246-259.

Vroomen, J., Keetels, M., Gelder, B. de, & Bertelson, P. (2004). Recalibration of temporal order perception by exposure to audiovisual asynchrony. Cognitive Brain Research, 22, 32-35.

Ward, R., & McClelland, J.L. (1989). Conjunctive search for one and two identical targets. Journal of Experimental Psychology: Human Perception and Performance, 15, 664-672.

Wertheim, A.H. (1981). Distraction in visual search. Soesterberg: TNO Institute for Perception (IZF 1981-7).

Wertheim, A.H., Hooge, I.T., Krikke, K., & Johnson, A. (2006). How important is lateral masking in visual search? Experimental Brain research, 170, 387-402.

Wertheimer, M. (1912). Experimentelle Studien über das Sehen von Bewegung. Zeitschrift für Psychologie, 61, 161-165.

Williams, L.G. (1966). The effect of target specification on objects fixated during visual search. Perception and Psychophysics, 1, 315-318.

Wolfe, J.M. (1994). Guided Search 2.0: A revised model of visual search. Psychonomic Bulletin and Review, 1, 202-238.

Wolfe, J.M., & Gancarz, G. (1996). Guided Search 3.0 Basic and Clinical Applications of Vision Science. Dordrecht, Netherlands: Kluwer Academic,189-192.

Wolfe, J.M., Cave, K.R., & Franzel, S.L. (1989). Guided search: an alternative to the feature integration model for visual search. Journal of Experimental Psychology: Human Perception and Performance, 15, 419-433.

Yantis, S., & Egeth, H.E. (1999). On the distinction between visual salience and stimulus-driven attentional capture. Journal of Experimental Psychology: Human Perception and Performance, 25, 661-676.

Yantis, S., & Hillstrom, A.P. (1994). Stimulus-driven attentional capture: evidence from equiluminant visual objects. Journal of Experimental Psychology: Human Perception and Performance, 20, 95-107.

Yantis, S., & Jonides, J. (1984). Abrupt visual onsets and selective attention: evidence from visual search. Journal of Experimental Psychology: Human Perception and Performance, 10, 601-621.

Yeshurun, Y. (2004). Isoluminant stimuli and red background attenuate the effects of transient spatial attention on temporal resolution. Vision Research, 44, 1375-1387.

Yeshurun, Y., & Carrasco, M. (1999). Spatial attention improves performance in spatial resolution tasks. Vision Research, 39, 293-306.

Yeshurun, Y., & Levy, L. (2003). Transient spatial attention degrades temporal resolution. Psychological Science, 14, 225-231.

Zampini, M., Shore, D.I., & Spence, C. (2003). Audiovisual temporal order judgments. Experimental Brain Research, 152, 198-210.

Zampini, M., Shore, D.I., & Spence, C. (2005). Audiovisual prior entry. Neuroscience Letters, 381, 217-222.

Zelinsky, G.J. (1996). Using eye saccades to assess the selectivity of search movements. Vision Research, 36, 2177-2187.

Zohary, E., & Hochstein, S. (1989). How serial is serial processing in vision? Perception, 18, 191-200.

Publications and presentations

Articles:

Van Schie, H.T., Wijers, A.A., Mars, R.B., Benjamins, J.S., & Stowe, L.A. (2005). Processing of visual semantic information to concrete words: temporal dynamics and neural mechanisms indicated by event-related brain potentials. Cognitive Neuropsychology, 22, 364-386.

Benjamins, J.S., Hooge, I.T.C., van der Smagt, M. J., & Verstraten, F. A. J. (2007). Disengaging attention stets the temporal limit of attentive tracking. Vision Research, 47, 1055-1059.

Benjamins, J. S., van der Smagt, M. J., & Verstraten, F. A. J. (in press). Matching auditory and visual signals: is sensory modality just another feature? Perception.

Benjamins. J.S., Hogendoorn, J.H.A. Hooge, I.T.C. & Verstraten, F.A.J. (submitted). Expect the unexpected: top-down suppression of attentional capture in time.

Benjamins, J.S., Hooge, I.T.C., Van Elst, J.C., Wertheim, L. & Verstraten, F.A.J. (submitted). Search time critically depends on irrelevant subset size in visual search.

Published Abstracts:

Benjamins, J.S., Van der Smagt, M.J., & Verstraten, F.A.J. (2004) Auditory induced visual illusion does not improve visual detection. Proceedings of the NWO Summer School: from stimulus to understanding in language and perception, [pp 21].

Van der Smagt, M.J., Benjamins, J.S., & Verstraten, F.A.J. (2004). Auditory induced visual illusion does not improve visual detection. Program No. 177.6. 2004 Online Abstract Viewer/Itinerary Planner. Washington, DC: Society for Neuroscience.

Benjamins, J. S., van der Smagt, M. J., & Verstraten, F. A. J. (2005). The upper temporal limit of attention-based motion perception is increased by an in-phase auditory stimulus. Journal of Vision, 5(8), 35a

Benjamins, J. S., van der Smagt, M. J., & Verstraten, F. A. J. (2005). The upper temporal limit of attention-based motion perception is not increased by an in-phase auditory stimulus. Proceedings of the 3rd Helmholtz Retreat, abstracts of posters [pp 5].

Benjamins, J. S., van der Smagt, M. J., & Verstraten, F. A. J. (2005). The temporal limit of binding sound and colour. Perception, 34, 29. (suppl.)

Benjamins, J. S., van der Smagt, M. J., & Verstraten, F. A. J. (2005). The temporal limit of binding sound and colour. Proceedings of the 10th NVP Winter Conference, [pp 53].

Benjamins, J. S., Hogendoorn H., Hooge, I.T.C., & Verstraten, F. A. J. (2007). Temporal properties of task-irrelevant events: attentional capture is not purely bottom-up. Journal of Vision, 7(9), 1082a

Posters:

Benjamins, J.S., Van der Smagt, M.J., & Verstraten, F.A.J. (2004, October) Auditory induced visual illusion does not improve visual detection. NWO Summer School: from stimulus to understanding in language and perception, Doorwerth, The Netherlands.

Benjamins, J. S., van der Smagt, M. J., & Verstraten, F. A. J. (2005). The upper temporal limit of attention-based motion perception is not increased by an in-phase auditory stimulus. Annual meeting of the Vision Sciences Society, Sarasota, United States of America.

Benjamins, J. S., van der Smagt, M. J., & Verstraten, F. A. J. (2005). The upper temporal limit of attention-based motion perception is not increased by an in-phase auditory stimulus.The 3rd Helmholtz Retreat, Bergen, The Netherlands.

Benjamins, J. S., van der Smagt, M. J., & Verstraten, F. A. J. (2005). The temporal limit of binding sound and colour. Annual Meeting of the European Conference on Visual Perception, A Coruna, Spain.

Benjamins, J. S., van der Smagt, M. J., & Verstraten, F. A. J. (2005). The temporal limit of binding sound and colour. 10th NVP Winter Conference, Egmond aan Zee, The Netherlands.

Benjamins, J. S., Hogendoorn H., Hooge, I.T.C., & Verstraten, F. A. J. (2007). Temporal properties of task-irrelevant events: attentional capture is not purely bottom-up. Annual meeting of the Vision Sciences Society, Sarasota, United States of America.

Benjamins, J.S., Hooge, I.T.C., Van Elst, J.C., Wertheim, L. & Verstraten, F.A.J.

(2007). Less is more: search time depends on irrelevant subset size. Annual Meeting of the European Conference on Visual Perception, Arezzo, Italy.

Oral presentations:

Benjamins, J.S. (2004). 'Parietal activations during judgments of brightness'. EuroGK meeting [host: Rijksuniversiteit Groningen/Frans Cornelissen] January 12, 2003, Groningen, The Netherlands.

Benjamins, J.S. (2004). 'Let op! Proef van de Dag'. Wetenweek 2004 [host: Universiteit Utrecht/Jesse Willemse] October 23, 2004, Utrecht, The Netherlands.

Benjamins, J.S. (2005). 'The temporal limits of binding sound and colour'. Helmholtz PhD-day [host: Psychonomie/Susan te Pas] December 6, 2005, Utrecht, The Netherlands.

Benjamins, J.S. (2006). 'Attentive tracking en temporele eigenschappen van aandacht.', Dag van de Perceptie, [host: TNO], October 13, 2006, Soesterberg, The Netherlands.

Benjamins, J.S. (2006). 'Attentive tracking and temporal properties of attention.', Helmholtz PhD-daydag [host: AMC Audiologie], November 8, 2006, Amsterdam, The Netherlands.

Benjamins, J.S. (2007). 'Temporele eigenschappen van aandacht', Colloquia Psychonomie, [host: Psychonomie], January 11, 2007, Utrecht, The Netherlands

Benjamins, J.S. (2007). 'Visual search does not always show set size effect', Heymans Colloquia Psychonomie, [host: Rijksuniversiteit Groningen, Psychologie], September 24, 2007, Groningen, The Netherlands.

Benjamins, J.S. (2007). 'Opletten!' Science Sunday 2007 workshop [host: Universiteit Utrecht] October 21, 2007, Utrecht, The Netherlands.

Benjamins, J.S. (2007). 'Matching auditory and visual signals: sensory modality might just be another feature' [host: Universiteit van Tilburg/Mirjam Keetels] Symposium Multisensory Perception, December 7, 2007, Tilburg, The Netherlands.

Dankwoord

Een proefschrift komt niet zomaar tot stand, maar is het resultaat van vier jaar werk. Een dankwoord is dan ook een dank aan een ieder die in die jaren heeft bijgedragen aan het tot stand komen van dit proefschrift, hetzij in de vorm van wetenschappelijke discussies hetzij voor de broodnodige gezelligheid en geouwehoer ter afwisseling van die wetenschap.

Ten eerste daarom dank aan mijn promotor, Frans Verstraten, voor de betrokkenheid bij en het gestelde vertrouwen in mijn onderzoek. De gelegenheid om dit onderzoek vervolgens wereldkundig te maken op internationale conferenties heb ik ook erg kunnen waarderen. Het is goed een promotor gehad te hebben die beseft hoe belangrijk het is je data te presenteren en kritisch te bediscussiëren met de deskundigen op je gebied.

Ook mijn copromotores, Maarten van der Smagt en Ignace Hooge, ben ik dankbaar voor hun steun, advies, commentaar en discussies in mijn onderzoek. Maarten, als ik van iemand heb geleerd hoe een manuscript aangescherpt en genuanceerd kan worden, dan ben jij dat geweest. Ik weet zeker dat ik in de toekomst nog vaak zal denken bij het opschrijven van een manuscript: “wat bedoel ik met deze zin?”. Verder weet ik zeker dat ik bij iedere te flauwe woordgrap of zinspeling, die ik hoor of zelf maak, zal denken: “wat doet Maarten hier?”.

Ignace, je bent als tweede copromotor betrokken geraakt bij mijn promotie, maar je bent zeker niet minder belangrijk geweest. Je hebt ooit gezegd een ‘katalysator met staart’ te willen zijn voor mij en dat ben je geweest. Een kwartiertje of een half uurtje met wat scherpe vragen en opmerkingen van jou was meestal genoeg om weer verder te kunnen met schrijven of programmeren. Dit heb ik erg gewaardeerd. Daarnaast zal ik dingen dingen als Need for Speed tot laat op de avond, de autoRAI, en een broodje integratie voor Ryota niet snel vergeten.

Niet alleen promotor en copromotores wil ik bedanken voor het tot stand komen van dit proefschrift, ook de verschillende mensen met wie ik samengewerkt heb in verschillende experimenten verdienen een woord van dank. Hinze Hogendoorn en ik hebben ons een week lang met alle mogelijk permutaties van eenzelfde paradigma die we konden bedenken suf gepilot, hetgeen geleid heeft tot het onderzoek in hoofdstuk 4 van dit proefschrift. Dank daarvoor, maar ook dankjewel voor nachtelijke ritten door Sarasota, de vele uren ‘klok kijken’ en af en toe ‘nog even’ naar België. Lex Wertheim wil ik bedanken voor het tijdens een kerstdiner zo enthousiast vertellen over een experiment dat ik het de volgende ochtend direct geprogrammeerd had. Met uitbreidingen heeft dit de reeks experimenten in hoofdstuk 5 opgeleverd. Lex, jouw enthousiasme en mooie verhalen over allerhande onderzoek, boeken, fotografie en wat al niet meer is aanstekelijk.

Ook Jacco van Elst is bij het uitvoeren van de experimenten in hoofdstuk 5 van groot belang geweest. Dank voor het meten van een deel van de psychofysica en de oogbewegingen. Daarnaast denk ik dat we, samen zittend achter MATLAB, beiden veel geleerd hebben, terwijl we de data analyseerden. Het was prettig met je samen te werken en de data door te ploegen.

Het beginnen met mijn promotieproject op een tijdelijke Mac in het lab, het daarna tijdelijk werken in Nijmegen vanwege asbestverwijdering, en vervolgens het verhuizen van 16.21 naar 17.07, naar 17.16 en uiteindelijk weer terug naar 17.07, heeft in een lange rits kamergenoten geresulteerd. Hubert Fonteijn, Meike Grol, Mark Rijpkema, Tanja Nijboer, Carinne Piekema, Tobias Borra, Katinka van der Kooij en Jeanette Lortije, dank voor jullie gezelligheid en slap geouwehoer in dit scala aan werkplekken van de afgelopen vier jaar.

Met veel plezier heb ik ook master- en bachelorstudenten begeleid in hun eindonderzoeken. Jasper Sens, Arlette Link, Mirte van Deursen, Chantal Groenewoud, Maartje de Vries en Marvin Meeng, ik hoop dat het jullie allen verder goed gaat in jullie carrières, hetzij binnen, hetzij buiten de wetenschap.

Ans Hörchner-Driehuis, Veronica Maassen, Janneke Kreek en Leonie Woudstra, dank voor alle hulp met allerhande paperassen, het printen van dit proefschrift voor de leescommissie, voor de drop en de altijd gezellige gesprekken.

Tanja Nijboer en Katinka van der Kooij, een van de eerste en mijn laatste kamergenoot, ik ben erg blij dat jullie mijn paranimfen willen zijn. Tanja, ik denk dat geen ander op de afdeling mij beter heeft leren kennen dan jij. Je bent een erg goede vriendin geworden in de afgelopen jaren en daar ben ik blij om. Ik denk dat er nog vele etentjes, avonden drinken en dansen, weekendjes Groningen, TV kijken via de chat, concerten of wat dan ook maar zullen volgen.

Katinka, naast slimme wetenschapper, ook nog eens getalenteerd schrijfster. Als er iemand is die altijd goed nadenkt over wat ze zegt en schijft, ben jij het wel. Ik heb met veel plezier een voorproefje van je verhalenbundel mogen lezen en ik kan niet wachten tot ik de hele bundel (uiteraard gesigneerd!) in mijn handen heb.

Verder dank aan alle anderen die op de 16^e en 17^e verdieping van het Van Unnik rondliepen voor gezellige avonden in de stad, heerlijke etentjes, goede conferenties, Helmholtz retraites, het samen organiseren van de NVP pubquiz en wat niet meer. Jelmer de Vries (dank ook voor het doorlezen van een deel van dit proefschrift), Titia Gebuis, Chris Paffen, Nisan Mol, Marieke Lansbergen, Floris Klumpers, Björn Vlaskamp, Susan te Pas, Ryota Kanai, Jan Souman, Mark Nieuwenstein, Peter Putman, Gudrun Nys, Erno Hermans, Thomas Carlson, Loes Koelewijn, Arielle Veenemans, Marnix Naber, Joukje Oosterman, Stefan van der Stigchel, Helen Anema, Marijn Struiksma, Jessie Bullens, Marjolein Kammers, Maartje de Goede, Daan Baas, Marga Veldhuizen, Sophie van Rijn, Bas Neggers, Mascha van 't Wout, Sander Zuidhoek, Dennis Schutter, Linda Ooijen, Ineke van der Ham en Lennart Verhagen, bedankt!

Ook iedereen uit Groningen - van tijdens mijn studie of al (veel) eerder - dank voor fantastische feesten ter ere van verjaardagen, promoties of voor zo maar gewoon eens even ergens afspreken. Rogier Mars, Carinne Piekema, Thomas Gladwin, Rianne Veenstra, Arjan Stoter, Marleen Groenier, Job Lindsen, Joost Driesens, René Scheeringa, Harmen Piekema, Joke Schaper, Dennis Westra, Hanneke Luth, Joyce Boucard, Sonja Tomaskovic, Eleonora Rossi, en Mbemba Jabbi, het was altijd fijn jullie weer te zien.

Last, but not least, bedank ik mijn familie. Zonder jullie zou de afgelopen periode een stuk zwaarder geweest zijn. Het is moeilijk in woorden uit te drukken hoe dankbaar ik jullie daarvoor ben. Misschien dat het het nog het beste is uit te drukken in hoe ik tegen jullie aankijk.

Joke, mijn lieve oudste zusje, de beste psycholoog die er is; jouw meelevendheid met je patiënten siert je in je werk. Als geen ander kun jij luisteren naar mensen en ze gerust stellen. Maar niet alleen in je werk, ook in je dagelijkse doen en laten ben je een schat! Ik ben trots op je en houd van je!

Hanna, mijn lieve jongste zusje, nu al een fantastisch goede advocaat en cultureel zeer bevlogen, in het verleden als ambassadrice van het Jongeren Uitburo en nu nog steeds, gewoon omdat je het leuk vindt. Jouw vastberadenheid en echt je eigen gang gaan vind ik schitterend. Ik houd van je en ben trots op je!

Als laatste, mijn ouders. Veel mensen zullen dit misschien zeggen van hun ouders; maar als er twee mensen zijn die hun hart op de juiste plaats hebben, dan zijn jullie het! Ik denk dat jullie dit op ons alle drie hebben overgedragen en ons hebben gevormd zoals we zijn. Ook mijn nieuwsgierigheid is door jullie altijd gestimuleerd. Het heeft mij zeker gevormd, juist ook in mijn onderzoek, en daar ben ik jullie dankbaar voor. Er is echt geen liever stel mensen dan jullie twee! Het is schitterend jullie te zien zitten in de werkkamer thuis. In volledige harmonie: pap achter de piano, mam achter de computer. Pap en Mam, heb ik vandaag al gezegd hoeveel ik van jullie houd?

Curriculum Vitae

Jeroen Sander Benjamins werd geboren op 21 maart 1980 - zoals het een echte Groninger betaamt - in Huize Tavenier te Groningen.

In 1998 haalde hij in Groningen zijn VWO diploma aan het Willem Lodewijk Gymnasium.

In datzelfde jaar is hij psychologie gaan studeren aan de Rijksuniversiteit Groningen. In september 2003, na een stageonderzoek bij Dr. Frans Cornelissen naar helderheidsperceptie en daaraan gekoppelde hersenactiviteit, ontving hij zijn doctoraal bul psychologie in de hoofdrichting Funktieleer en nevenrichting Neuro/biopsychologie

Na het behalen van zijn bul begon hij in oktober 2003 met zijn promotieonderzoek bij Prof. dr. Frans Verstraten en co-promotores Dr. Maarten van der Smagt en Dr. Ignace Hooge. Momenteel werkt Jeroen aan de toepassing van oogbewegingsmetingen in reclame-uitingen bij Intomart Gfk te Hilversum.

SR Research

EyeLink®

Complete
Eye Tracking
Solutions