Chapter 1
Introduction
INTRODUCTION

Many of our perceptual tasks are based on processing only a small amount of the sensory information that is available at a given time. In order to perform tasks such as reading and visual search for objects, we need a system that helps us to focus on certain stimuli whilst ignoring the competing information overflow. Attention has been proposed to be the mechanism that enforces this selection (James, 1890). Furthermore, attentional selection has been assumed to be necessary in order for a stimulus to become consciously perceived (Cohen, Cavanagh, Chun, & Nakayama, 2012), even though this has been disputed (Lamme, 2003). Importantly however, treating attention in a binary fashion, and considering its effects in terms of whether it is active or not, ignores the strong temporal dimension that attentional selection is assumed to have. The strength of a perceptual object may evolve gradually, as attention directed to the object becomes more effective. It is necessary to understand the temporal dynamics of the process: how fast does the selection develop and how long will it last. The time course of attentional selection is the topic of this thesis.

The thesis focuses on the time course of attentional selection as based on two types of signals in particular. First, we select visual information based on external cues in the environment, such as when a salient colored object or a flash of light draws our attention. In addition, attentional selection can occur via an internally represented template. Such selection is at stake for example when you are searching for your keys, and presumably occurs on the basis of some mental picture of the keys held in memory. The former type of selection is referred to as bottom-up or exogenous and the latter as top-down or endogenous (Carrasco, 2011). Whether attentional selection occurs exogenously or endogenously is thus determined by the attention-triggering signal, i.e. either an external, bottom-up cue or an internal, top-down representation. Furthermore, an exogenous signal that focuses attention to a certain object is by definition spatial in nature: A flashing light or a brightly colored object in a homogenous background inherently has a certain location to which attention can then be drawn. In contrast, endogenous signals are not necessarily spatial in nature, but can be feature-based, for example specifying the color and shape of your keys without knowing where they are.

Crucially, the two types of attentional mechanisms are also suggested to operate at identifiable and mutually separable time courses. For example, Nakayama and Mackeben (1989) showed that exogenous attention as induced by a spatial cue was rather fast and short-lived, whereas endogenous attention as induced by a description of the target features appeared to be slower but sustained for a longer period of time. Although the idea of rapid and transient exogenous attention, versus slow and sustained endogenous attention has been widely adopted, it has remained elusive what type of
attentional operations or other factors underlie these temporal profiles. The aim of this thesis is to further clarify how selection occurs in time, through external signals or internal representations. In the current chapter I will review existing studies on the time course of attention, in relation to the research conducted for this thesis.

**The time course of attentional selection based on external signals**

**Evidence from the spatial cueing task**

Our experience tells us that visually salient events capture attention automatically and abruptly. An often used paradigm in the studies of attentional allocation is the spatial cueing task (Posner, 1980). In one version of this paradigm, a cueing signal is first presented, potentially indicating where the target will be shown. Then, after some time from the cue onset (referred here as the stimulus onset asynchrony, SOA), a target appears and observers are asked to either detect its occurrence or to discriminate some of its features, such as the orientation of a line. The effects of spatially specific attention are then assumed to be reflected in response speed or accuracy changes as a function of cueing. The primary focus in the following discussion will be on the effects of external, location-specific cues on perceptual discrimination, in studies that have systematically varied the time of cueing.

By measuring perceptual discrimination accuracy for targets at the cued location, a number of studies have found evidence for what appears to be the standard time course of exogenous selection (Chastain, Cheal, & Lyon, 1996; Chastain & Cheal, 1997, 1998; Cheal & Lyon, 1991; Kristjánsson, Mackeben, & Nakayama, 2001; Müller & Findlay, 1988; Müller & Rabbitt, 1989; Nakayama & Mackeben, 1989; Mackeben & Nakayama, 1993; see B.-T. Lee & McPeek, 2013 for similar findings in macaque monkey). The consistent finding from these studies is that target discrimination accuracy grows rapidly to an optimal level when the SOA is increased, up to about 100-200 ms from the cue onset. However, after this enhancement period, performance is often found to decline again. The resulting cued performance pattern, including the rapid enhancement and subsequent decline, has been proposed to represent a *transient attention* mechanism (Nakayama & Mackeben, 1989; see also Weichselgartner & Sperling, 1987). The fact that this transient attention performance function was found to follow a salient location cue and develop so rapidly differentiated it initially from a slower time course of sustained type of attention, that occurred as a response to a centrally presented symbolic cue (Cheal & Lyon, 1991; Müller & Findlay, 1988; Müller & Rabbitt, 1989). Similarly, salient stimuli have been found to have
INTRODUCTION

A rapid and short-lived influence on saccadic target selection, assumed to reflect the transient attention mechanism (Van Zoest, Donk, & Theeuwes, 2004).

Nakayama and Mackeben (1989) originally proposed that transient attention reflects an automatic component of attention, as they found a similar pattern irrespective of top-down knowledge about the target location, that is, in a condition in which the cue and target position was fixed. However, finding that the cued function does not change by top-down information does not logically imply that there is no top-down control on the transient attention time course. Importantly, in most studies of transient attention the cue fully predicts the target location, so that there is no reason why observers would not deliberately use this information (but see Müller & Findlay, 1988; Müller & Rabbitt, 1989 for transient attention studies that used 50% valid cues). Even in the cases where the cue is not fully predictive, it may still be informative in that it is contingent with the target settings in terms of sudden onset and possible location (Ansorge & Heumann, 2003). Furthermore, some recent studies have shown that the peak of transient attentional enhancement can be modulated by spatial and temporal consistencies between trials (Kristjánsson, et al., 2001; Kristjánsson, Eyjófsdóttir, Jónsdóttir, & Arnkelsson, 2010), suggesting that implicit type of learning can change the outcome of cueing. In short, there is evidence that the transient performance time course reflects mainly automatic attentional selection that is initiated by direct, external cues. Nevertheless, the response may be modulated by top-down control and learning. An interesting puzzle that remains is why attention is consistently found to be transient even when the cue is fully predictive and observers are thus fully aware of the target location (Chastain, et al., 1996; Chastain & Cheal, 1997; Cheal & Lyon, 1991; Kristjánsson, et al., 2001; Mackeben & Nakayama, 1993; Nakayama & Mackeben, 1989). With such predictive cues, the best strategy would be to sustain focus at the cued location until the target appears. In fact, in another paradigm, a similar transient enhancement of attentional selection has been proposed to occur even though the target location is always the same from trial to trial, as we will see next.

Evidence from rapid serial visual presentation

The time course of attentional selection has also been studied in rapid serial visual presentation (RSVP) tasks. In this type of task, a stream of stimuli is presented sequentially, typically at a single, central location. The task is to report one or more predefined items, while ignoring the other, irrelevant items. An often replicated finding from the RSVP task is that when two targets are presented, the identification of the second target follows a U-shaped time course time-locked to the first target. This is known
as the attentional blink (Raymond, Shapiro, & Arnell, 1992; Chun & Potter, 1995).

Specifically, when performance for the second target is plotted as a function of the first target onset, there appears to be an initial period of higher target accuracy for about 100 ms, followed by a decline between 200-500 ms, and a subsequent improvement back to the initial level of performance. Even though the performance time course is usually sampled only sparsely (every 100 ms or less), the initial peak at about 100 ms SOA has been proposed to represent a transient attentional enhancement (Bowman & Wyble, 2007; Nieuwenhuis, Gilzenrat, Holmes, & Cohen, 2005; Olivers & Meeter, 2008; Wyble, Bowman, & Nieuwenstein, 2009).

There are different possibilities why the first target in RSVP stream may attract transient attention. First, in some studies the first target was salient relative to the rest of the items, as it was presented in an unique color (Nieuwenstein, Chun, van der Lubbe, & Hooge, 2005; Raymond, et al., 1992), or highlighted by an outline square (Weichselgartner & Sperling, 1987). By contrast, in other studies the target could be differentiated from the distractors by category, as when letters had to be reported amongst digits (Chun & Potter, 1995; Wyble, Bowman, & Potter, 2009). Targets defined by category produce less of a salient signal, but may still trigger an automatic attentional response on the basis of their match to the top-down task settings (Folk, Remington, & Johnston, 1992; Nieuwenstein, et al., 2005; Wyble, Bowman, & Potter, 2009). Note that such an attentional response (also referred to as contingent capture; Folk, et al., 1992) is not purely exogenous, as it depends on the goals of the observer. It is neither purely endogenous as it is not initiated by the observer, but by the target-matching stimulus. Furthermore, the time course of the task-contingent enhancement has also been shown to be rapid, occurring already as fast as 50-100 ms after the target (Folk, et al., 1992), and is also transient, declining at longer times (Wyble, Bowman, & Potter, 2009). Therefore, through a salient signal, or signal matching with the task settings, the first target in the RSVP stream may enhance attention quite similar as in spatial cueing.

There is however some ambiguity whether the attentional enhancement in RSVP is spatially specific, i.e. restricted to the target location (Bowman & Wyble, 2007; Nieuwenhuis, et al., 2005; Wyble, Potter, Bowman, & Nieuwenstein, 2011), as is assumed in spatial cueing. Other important aspects that may differentiate between attentional selection as found in the spatial cueing and RSVP tasks, is that in the latter the first target needs to be encoded and reported, and distractors interleave the two targets of the sequence. In fact, the performance suppression that is found at intermediate times has been assumed to be due to the encoding of the first target (Bowman & Wyble, 2007; Chun & Potter, 1995; Potter, Staub, & O’Connor, 2002) or because of the intervening distractors
(Di Lollo, Kawahara, Ghorashi, & Enns, 2005; Olivers & Meeter, 2008). Spatial cueing does not typically involve either of these factors, so that it remains elusive to what extend do the performance time courses in the two tasks correspond to the same attentional processes. The current thesis bridges some of this gap, as it uses spatial cueing tasks that come closer to the RSVP paradigm, in which cues and targets are presented centrally, at one location.

The time course of attentional selection has been investigated in modified RSVP tasks that better resemble cueing (Nieuwenstein, Van der Burg, Theeuwes, Wyble, & Potter, 2009; Chapter 4; see also Nieuwenstein, Potter, & Theeuwes, 2009). Specifically, they presented only two targets of which the second one was masked, but no distractors in between. In addition, some of the experiments only required a response to the second target, giving the first target a cue-like function. These studies have found a performance pattern similar to that in the more typical RSVP task, with initial sparing, decline, and recovery. This suggests that the time course may not be dependent on the intervening distractors. In addition, neither did performance seem to be influenced by the response and encoding requirements of the first target. However, the experiments of the current Chapter 4 show that performance pattern in the truncated RSVP task can be modified by both lower and higher level properties of the cue and the target. Specifically, Experiment 4 of Chapter 4 found that the early part (< 83 ms) of the letter discrimination time course was affected by cue salience, whereas performance was later modified by the shapes of the cue and the target. When the cues and targets were letters, as is often the case in the RSVP tasks, performance was suppressed at the intermediate SOAs (between 83-233 ms). It was further shown that this suppression may result from confusion or swap between the cue and target identities. Note that cueing studies typically use clearly separable cue and target stimuli whose identities are less likely to be confused.

Other evidence for a transient attention time course

Transient attention type of performance has also been described in studies that have used other measures than target discrimination accuracy or speed (Bachmann, 1988; Scharlau, Ansorge, & Horstmann, 2006; Suzuki & Cavanagh, 1997). Instead, these studies have looked at how contrast estimates, temporal order judgments, and perceptual repulsion effects are influenced by external location cues at variable SOAs. For example, Suzuki and Cavanagh (1997) found that a cue made a line to appear displaced from its location, and that the magnitude of this perceptual repulsion effect increased and declined as a function of cueing SOA, similar to the transient attention function. By using a temporal
order judgment task, Scharlau et al. (2006) showed that a spatial prime speeded stimulus perception with a time course that corresponded to transient attention. In the task of Bachmann, contrast estimation of a letter was improved by a preceding letter, again with a function that peaked at about 110 ms and declined at later times. Whereas Suzuki and Cavanagh, and Scharlau et al. interpreted the observed transient performance enhancement in their tasks as a spatially specific attentional effect, Bachmann suggested that the rapid improvement in his contrast estimation task was due to a more non-specific thalamic facilitation.

Figure 1. Performance time course as a function of exogenous cueing. Separate curves indicate results from different conditions (A), different experiments (D), and a different plotting method (B). Adapted from A) K. Nakayama & M. Mackeben, 1989, Sustained and transient components of focal visual attention, Vision Research, 29, 11, p. 1638 with permission from Elsevier; B) E. Weichselgartner & G. Sperling, 1987, Dynamics of automatic and controlled visual attention, Science, 238, p. 779 with permission from AAAS; C) S. Suzuki & P. Cavanagh, 1997, Focused attention distorts visual space: An attentional repulsion effect, Journal of Experimental Psychology: Human Perception and Performance, 25, 2, p. 452 with permission from APA; and D) I. Scharlau, U. Ansorge, & G. Horstmann, 2006, Latency facilitation in temporal-order judgments: Time course of facilitation as a function of judgment type, Acta Psychologica, 122, p. 152, with permission from Elsevier.
INTRODUCTION

The transient performance pattern has thus been reported in a variety of tasks that involve target cueing, most prominent being the spatial cueing and RSVP paradigms. Some examples from these different task types are presented in Figure 1, all suggesting a performance improvement at about 100-200 ms, and a subsequent decline. The current thesis clarifies the factors that explain the transient attention pattern in spatial cueing.

Which attentional operations could lead to transient pattern of performance?

An important objective is to resolve which attentional processes specifically cause the characteristic rapid enhancement and subsequent decline in performance, as a result of an external cueing signal. At least three components of attention has been identified that can contribute to performance in the spatial cueing task (Posner & Boies, 1971; Posner & Petersen, 1990; Raz & Buhle, 2006), and one or more of these could underlie the transient attention time course. First, spatial reorienting or shifting is thought to be necessary in spatial cueing to move the focus from fixation to the cued location. Second, the occurrence of the cue may evoke alerting that enhances response speed and perception in a spatially non-specific fashion. Third, selectivity is needed when a target is presented with competing distractor items. This has often been the case when the effects of cueing on perceptual discrimination accuracy have been examined. Altogether these three components are most readily identifiable in spatial cueing, whereas not all of them operate in the RSVP task. Spatial reorienting is one such a factor.

Spatial reorienting

Some of the studies that have examined the performance time course by exogenous cueing have assumed that the transience of attention is at least partially determined by the time that it takes to reorient attention, i.e. to shift attention from the fixation to the cued location, and away again (Cheal & Lyon, 1991; Müller & Findlay, 1988; Müller & Rabbitt, 1989). In the classic RT version of the spatial cueing paradigm, speeded target detection was measured as a function of exogenous cues that typically indicate the target location in only 50% of the cases. A common finding is that at relatively short SOAs, typically about 100 ms from the cue, RTs are improved at the cued location. However, when the cueing time is increased further, beyond 200 ms, RTs at the cued location become slower compared to other, uncued locations, a phenomenon that has been termed inhibition of return (IOR; Klein, 2000; Posner & Cohen, 1984; Posner, Rafal, Choate, & Vaughan, 1985). This biphasic facilitatory-inhibitory pattern has been typically explained by reorienting attention to the cued location and subsequently inhibiting or withdrawing...
from it. The fact that IOR and transient attention share a similar time course and both are observed in the spatial cueing task, suggest that they may be one and the same thing.

However, the conditions necessary for finding IOR also differ in important aspects from those of transient attention studies. Namely, in IOR studies RT is typically used as the outcome measure (see however Handy, Jha, & Mangun, 1999 for an example of IOR effect on perception), often in conditions in which a single target is presented for an unlimited duration. Cue is valid in 50% of the trials, and it appears that IOR cannot be observed if the cue is predictive (Wright & Richard, 2000). In contrast, transient attention studies have measured performance accuracy for targets that are presented shortly and backward masked. The target is presented together with irrelevant distractor items, and in the majority of the studies the cue was always valid (with the exception of studies by Müller & Findlay, 1988; and Müller & Rabbitt, 1989). It seems then that the conditions that are common for observing transient attention and IOR are fewer than those that are not.

Nakayama and Mackebe (1989) suggested that the transient pattern was not related to spatial reorienting, as they presumed that observers could not shift their attention back and forth as rapidly as the enhancement and decline of the performance function indicated (which is strange given the earlier IOR findings). In addition, the proposal that transient attention occurs in RSVP (e.g. Weichselgartner & Sperling, 1987) would also require the assumption that spatial shifting plays a little role in causing the later decline of performance, since in RSVP all items are typically presented sequentially at a single, central location. In fact, Nakayama and Mackebe (1989) also found preliminary evidence for this by showing the typical transient attention pattern for locations close to fixation, and for a peripheral but constant location. However, as observers presumably started from central fixation, there was still some shifting involved in these experiments (plus the effects were shown for two observers only).

Chapters 2 and 3 of the present thesis provide compelling evidence that spatial reorienting is not necessary for observing the transient attention performance pattern. In Chapter 2 discrimination performance was investigated in a task that was modified from the spatial cueing experiments of Nakayama and Mackebe (1989). Importantly, spatial shifting was made unnecessary by presenting the cue and target always at central fixation. The result that was replicated in a series of experiments was that cueing first improved performance which then declined, as is characteristic of transient attention. In Chapter 3 similar results were found in a spatial cueing task, in which the cue and target locations were varied. These experiments thus consistently show that cueing results in the standard
transient attention time course irrespective of whether spatial reorienting is part of the task or not.

*Non-specific alerting*

A second factor that should be considered when explaining the cued performance time course is spatially non-specific alerting or preparation. The occurrence of a spatially non-specific signal is known to improve performance with increasing cue-target SOA (Los & Schut, 2008; Niemi & Näätänen, 1981). Alerting effects are found to improve up to some optimal level, after which RTs usually stabilize, although in some cases the benefits may start vanishing again (Müller-Gethmann, Ulrich, & Rinkenaue, 2003). The time at which optimal performance is reached is thought to be determined by the effective preparation period which can be further influenced by expectancy about when the target may occur (Niemi & Näätänen, 1981). Such effects have been found for cues and targets in different sensory modalities, supporting the idea that the alerting response is non-specific in nature. Furthermore, the performance improvement is typically observed in RTs, and the alerting effect has been thought to reflect speeded motor responses or response preparation.

Even though it is often assumed that non-specific alerting exerts its effects through facilitating response-related processes, some studies have found evidence for non-specific alerting effects on perception (Jepma, Wagenmakers, Band, & Nieuwenhuis, 2009; Matthias, et al., 2010; Rolke & Hofmann, 2007). For instance, Jepma and colleagues (2009) found that an auditory stimulus improved letter discrimination by facilitating target encoding, as suggested by results from diffusion model analysis and ERP measurements. Matthias et al. (2010) examined the effects of a non-specific visual stimulus on visual target processing. In their task a large square cue surrounded the whole letter target display. Observers were asked to report as many letters as possible in an unspeeded manner. The authors then estimated three parameters of the theory of visual attention (Bundesen, 1990) in the cued and uncued condition, at several cue-target SOAs. Specifically, visual perceptual processing speed, visual short term memory processing capacity, and spatial bias of attentional weighting were examined. The presence of the global, spatially non-specific cue was found to improve the perceptual effectiveness (speed of processing), but not the visual short term memory capacity. The spatial attentional weights appeared to be biased towards the left visual field when the alerting cue was present. Perceptual processing speed was improved by alerting cues especially at the 80 ms SOA, whereas at longer SOAs both cued and uncued conditions improved this
component. Based on these results it was suggested that a non-specific alerting cue initially improves perception in a rapid and transient way, whereas intrinsic, top-down controlled alertness has a slower but sustained effect that benefits perception at later cueing times. Possibly, a non-specific alerting cue could then improve performance at similar latencies as a direct location cue. However, as the study did not include a condition with spatially specific cues, the congruence between the temporal effects of alerting and spatial cueing remains elusive.

The temporal dynamics of general cueing effects can also be inferred from studies that used symbolic cues to investigate endogenous attention (Cheal & Lyon, 1991; Müller & Findlay, 1988; Müller & Rabbitt, 1989). These studies have shown that target discrimination is improved up to about 300 ms and is then further sustained when the cue is presented centrally. Even though such a cue provides information about the target location, it should also act as a strong general alerting event. Since in these studies no transient component was found, we must conclude that general alerting cannot explain the rapid and transient attention time course that is found for spatial cues.

To directly investigate the effects of alerting, Chapter 2 tested whether the transient attention time course, as was found for spatially specific cues, could also be produced by spatially non-specific, global cues. In Experiment 3 a target was always presented at central fixation, preceded by either a local or global cue. The local cue was shown at the target location, whereas the global cue surrounded the whole stimulus display that included targets and distractors. Note that observers always knew the exact (central) target location in both conditions. The time course for spatially specific cues had the typical rapidly increasing and then declining pattern, whereas when the cue was location non-specific, accuracy was overall much lower and did not vary as a function of cueing time. Therefore, these results support the idea that transient attention requires a spatially specific external signal, whereas general alerting has dissociable effects.

Moreover, Chapters 2 and 3 show that the transient performance time course can be observed only in conditions in which the target is accompanied by irrelevant distractor items. When targets were preceded by the same cues, but when no distractors were present, performance increased rapidly and was then sustained or even improved further at SOAs beyond 100 ms. This suggests that the transient attention time course is specific to the conditions in which selectivity is needed to overcome competition by distractors. In contrast, the effects of alerting have typically been observed for single targets, presented without distractors. In other words, the non-specific alerting effects as induced by cueing do not require selection. Therefore, while alerting or preparation
INTRODUCTION

effects may contribute overall to cued performance, they appear not to contribute to the transient component of that performance. The role of selectivity in explaining the temporal effects by external cues will be discussed next.

Selectivity

Although it has been frequently debated whether competition by distractors, and hence selectivity, is needed for attention to reveal its transient component (Carrasco, Penpeci-Talgar, & Eckstein, 2000; Dosher & Lu, 2000; Shiu & Pashler, 1994, 1995; Yeshurun & Carrasco, 1998), it has been rarely investigated. So far, the majority of spatial cueing studies that found transient attention with fully predictive cues, have used displays that always contained distractors, but without manipulating their presence (Chastain, et al., 1996; Chastain & Cheal, 1997; Cheal & Lyon, 1991; Kristjánsson, et al., 2001; Mackeben & Nakayama, 1993; Nakayama & Mackeben, 1989). Müller and Findlay (1988) have been the only ones investigating the effect of the presence of non-targets while also manipulating cue-target SOA within the range that typically covers the transient attention function. They found that when the cue validly predicted the target location, performance rapidly improved but then declined, resembling the transient attention time course. Interestingly, the time course was similar, and thus transient both when three distractors accompanied the target and when the target was presented alone. However, even in the target only condition, the target was not exactly alone. The displays always consisted of placeholders at all possible stimulus positions. These in themselves may have acted as competing signals. In addition, the cue in their study had only 50% validity, so observers may have been discouraged to maintain their attention at the cued location, leading to IOR.

Another study that investigated the effect of distractors on spatial cueing found that performance was worse for letter targets presented with distractors, as compared to targets presented alone (Bachmann, Mäger, Sarv, Kahusk, & Turner, 1999). However, the effect of distractor presence was inconsistent, and could either increase or decrease over time. In this study the longest cue-target SOA used was 320 ms and the results showed no asymptote or downturn of performance within the measured time range. Possibly, the temporal cueing effects were different from the typical transient attention function as letters were used as cues and targets. Indeed, Chapter 4 of this thesis suggests that letter stimuli lead to different cueing effects than less meaningful stimuli like vernier bars.

Chapters 2 and 3 represent the first studies that measure the time course of cueing as a function of the presence of competing distractor signals. These experiments show consistently that the presence of distractors modifies the time course of cued
attention. Performance improved rapidly by cueing, irrespective of the distractor condition. However, after about 100 ms performance was sustained at the asymptotic level, or sometimes even improved further if the target was presented alone. By contrast, if the target was presented with distractors, performance dropped significantly at the longer SOAs. These findings thus strongly suggest that it is the competition for selection that underlies the transient component of the attention time course.

**Does the type of competition affect the transience of attention?**

Given that the presence of competition results in the transient attention time course, is this then dependent on the type of the competition? Visual discrimination performance is known to suffer from competition at lower sensory levels, as well as at higher levels of object identification. For example, the literature on *crowding* has demonstrated that target discrimination can suffer from distractors that perceptually group with the target (Kooi, Toet, Tripathy, & Levi, 1994; Manassi, Sayim, & Herzog, 2012; Saarela, Sayim, Westheimer, & Herzog, 2009) or that are confused with the target identity or location (Dakin, Bex, Cass, & Watt, 2009; Greenwood, Bex, & Dakin, 2010; Strasburger & Malania, 2013). It is then interesting to look at whether some of these different competition types would modulate the time course of exogenously cued attention in a distinct manner, possibly explaining the transient pattern further.

Only few cueing studies have investigated the effects of different types of competition on the cued time course of attention. Chastain and colleagues (Chastain, et al., 1996; Chastain & Cheal, 1997) studied how the cueing function is shaped by the response compatibility between target and distractors (B. A. Eriksen & Eriksen, 1974). They found that for a 200 ms period following the cue, target discrimination performance was impaired by response incompatible distractors. The competition by response incompatibility was however found to be *additive* with the effects of cueing time, i.e. the observed performance time course had a transient shape both when distractors were compatible with the target and when they were not. Chapter 3 of the present thesis expands these findings and shows that the target-distractor compatibility does not interact with the effects of cueing, even when longer SOAs (up to 400 ms) are included that better covers the typical decaying part of the transient attention function. Together, these results provide evidence that response compatibility does not affect the time course of selection, nor explain the general effect of competition at the later SOAs. Thus, competition at the level of identity (i.e. the identity on which the response is based) is not necessary for attention to become transient.
Alternatively, the transient attention pattern may be caused by competition by feature similarity and grouping. For example, feature similarity between target and distractors can impair both detection and discrimination tasks (Baylis & Driver, 1992; Kooi, et al., 1994; Malania, Herzog, & Westheimer, 2007). Chapter 3 therefore also investigated whether competition by feature similarity modifies the time course of cued attention. Experiment 2 showed that when the target had the same polarity as distractors, performance was significantly lower than if these two were dissimilar, consistent with earlier studies that did not manipulate time. However, like the effects of distractor compatibility, feature similarity impaired performance to a similar extent throughout the cueing intervals. In other words, it did not modify the transient attention time course. Therefore, neither did competition at feature level of processing appear to be necessary for the cueing response to decline.

It is possible that the transient attentional response reflects selection among competition in general, regardless of the type of competition. This would also be compatible with the fact that Müller and Rabbitt (1989) found transient attention for T-shaped targets regardless of whether the competing elements were crosses or placeholder boxes. Still, one central factor that is known to increase competition in visual perception, and that may thus also explain why attention becomes transient in a cueing task, is the spatial proximity of the discrimination items (Bouma, 1970; B. A. Eriksen & Eriksen, 1974; Whitney & Levi, 2011). Chapter 3 also examined whether competition in space may be the determining factor behind the downturn of the cued attentional response. The results were mixed, as a between-subjects manipulation suggested that the decline occurred only for closer target-distractor distances, whereas direct comparisons within observers did not show an effect of spatial distance. There too a transient cueing pattern occurred similarly for targets that were flanked by distractors both close (1° of visual angle) or further away (up to 10°). These results thus lead to the conclusion that the spatial proximity of distractors is not crucial determining factor for the transient function to emerge. Altogether, the results of Chapter 3 suggest that the transient attention pattern reflects selectivity in general, irrespective of the level of competition.

**Interim conclusions: Attentional selection by external signals has a transient time course**

The available literature agrees that visual discrimination performance is rapidly (within about 100 to 200 ms) enhanced by external spatial cues, and may then decline afterwards. The evidence provided in this thesis suggests that this enhancement turns into a decline.
due to the presence of competition by irrelevant visual information, indicating that the transient pattern relates to the selectivity component of attention. The findings in the current thesis also show that this competition may play out at different levels, with different types of competition each leading to a similar time course. What does affect the time course are interactions between the cue and the target both at lower sensory and higher object level, as shown in Chapter 4.

The time course of attentional selection based on internal representations

In addition to the rapid, externally driven attentional response, a slower internally controlled, endogenous attentional process has been assumed by many attention theories (Müller & Rabbitt, 1989; Nakayama & Mackeben, 1989; Van Zoest, et al., 2004; Weichselgartner & Sperling, 1987). The second part of this thesis focuses on this latter type of attentional selection and specifically its time course. Two types of cueing approaches have been used to investigate the temporal dynamics of endogenous selection. First, the spatial cueing task that was described earlier has been used to examine how attention can be focused based on top-down information about the location. Second, selection has been studied as a function of feature or identity cues, to investigate how attention can be focused based on non-spatial information about the target. This is a common task in everyday life as we often select objects based on knowledge about their appearance. Chapters 5 and 6 present experiments that were conducted to study such selection that is based on internal representation of the target. Before discussing these and other results on the topic, spatial cueing studies about the time course of endogenous selection will be briefly covered.

Selection based on endogenous spatial cues

Eriksen and Collins (1969) were the first to investigate the time course of endogenous spatial selection by using the cueing procedure. They used two types of centrally presented cues which indicated the target location at variable SOAs. The cue could be either a number referring to one of predefined screen positions, or an arrow pointing to the target location. The arrow cue could thus be relatively directly used to focus attention, and it also resulted in higher overall performance than the number cue, which would need to be first translated to a corresponding location. However, the cueing time course was very similar in both cases, as letter target identification improved up to about 200 ms, after which there was no further observable change. However, the target displays were
not backward masked which may have allowed for identification based on the extra time provided by a visual afterimage, possibly confounding the estimates of selection. Later spatial cueing studies have used masked targets when investigating the timing aspects of endogenous selection (Cheal & Lyon, 1991; Müller & Findlay, 1988; Müller & Rabbitt, 1989). These studies too used arrow cues that preceded the target at the center of the stimulus display. Now performance was found to increase at a somewhat slower pace, taking about 300-400 ms to reach asymptote when displays were masked. However, once the top-down selection of a location became optimal, this level was sustained up to the longer cueing intervals, similar to what was suggested by the study of Eriksen and Collins (1969). The slower developing and sustained time course then differentiated it further from the transient time course of selection, as induced by external cues.

Note that external, spatially-specific cues may also evoke an endogenous attentional response, as discussed in the previous section. In conditions where the cue is predictive of the target location, it should result in selection not only by evoking bottom-up activation, but also because the observer can voluntarily attend to its location. Indeed, as was found in Chapters 2 and 3, cued performance was sustained from about 200 ms onwards, as long as there were no competing distractors.

Selection based on a target template

Instead of selecting on the basis of a location, indicated by a symbolic cue, selection can also be based on a feature, as indicated by some form of instruction (e.g. “look for the red object”). This instruction is assumed to be translated into an internal visual representation — the template. An experimental paradigm in which attentional selection is based on such an internal representation of the target identity is visual search. In visual search, a target is embedded among distractor items and typically the task of the observer is to search through the stimuli and indicate whether a predefined target is present or not, or report one of its features. Importantly, according to most theories of visual attention, visual search is performed based on a target template, which is furthermore held in working memory (Bundesen, 1990; Desimone & Duncan, 1995; Duncan & Humphreys, 1989; Wolfe, 1994). Such a template is assumed to be activated depending on the goals of the observer, and used to bias the competition between relevant and irrelevant information in the visual input. Conversely, visual search performance provides a means to investigate the time course of template-based selection.

Recent studies have examined the time course of template-based selection, by cueing search targets with feature or identity cues. Such cues may either be visual
(showing the target feature or image of the target), or categorical by nature (providing a verbal description of the target). The effects of visual cues will be discussed first.

Wolfe, Horowitz, Kenner, Hyle, and Vasan (2004) cued a search target by presenting a picture of it in the center of the display, at variable SOAs before the appearance of the search array. The task of the participants was to respond as fast as possible based on whether the cued target was present or absent. The results showed that response RTs improved rapidly and reached asymptotic levels as fast as 200 ms after the cue. Similar results were found by Vickery, King, and Jiang (2005), who also measured search RTs as a function of a picture cue. In their Experiment 4 the SOA was varied from 200 to 1000 ms, and selection of the cued target was found to be optimal already at the shortest time that was measured. These results suggest that attentional selection on the basis of a template may take only a maximum of 200 ms, and likely even less. This would come very close to the time estimate of selection based on external signals, and would go against the accounts of temporal attention that assume that exogenous and endogenous selection operate at clearly separable time courses (Müller & Rabbitt, 1989; Nakayama & Mackebe, 1989; Van Zoest, et al., 2004; Weichselgartner & Sperling, 1987).

However, other cued visual search studies have found selection times of about 400-500 ms (Meyers & Rhoades, 1978; Schmidt & Zelinsky, 2011; Chapters 5 and 6 of the present thesis), or even as long as 800 ms (Knapp & Abrams, 2012). The estimates of the time course of attentional selection as based on visually cued target templates thus vary considerably.

An important explaining factor may be the differences in the methods used, notably the dependent measure. Most of these studies have examined RTs (Knapp & Abrams, 2012; Meyers & Rhoades, 1978; Vickery, et al., 2005; Wolfe, et al., 2004), but as we argue in Chapter 5, RTs and even the slopes of the RT functions may not reflect the time course of selection that well (see also Wolfe, et al., 2004). Specifically, the basic RTs can represent different aspects of task performance, such as response preparation and execution, possibly camouflaging the effects of attentional selection. This is often circumvented by manipulating the set size, and calculating the RT slopes, that then better represent the actual time that is spent in processing each item. However, also the slopes appear to provide an inaccurate estimate for selection, when measured as a function of cueing time, because the unlimited response time allows for responding only after the template has been fully set up and implemented. In other words, observer may first wait until they have loaded their template, and only then start the search. If so, SOA will have little effect on search slopes. Indeed, the results of Chapter 5 support these claims as RT
INTRODUCTION

slopes were found to be unchanged by the cueing time. In contrast, slopes in an otherwise identical task using accuracy with limited-duration target displays as a dependent measure, were found to be gradually enhanced with time. This time course was replicated in Chapter 6.

Chapters 5 and 6 further differed from previous cueing experiment as they did not measure the effects of visually present cues, but on the basis of cues stored in memory. Two possible target colors were first memorized, after which they disappeared. A neutrally colored (black) postcue then indicated the relevant template for the search. The task was to report the location of the cued target within the subsequent search display. Importantly, the search display was presented for a limited duration only, after which it was covered by a backward mask. In this way, performance accuracy instead of RT could be measured. In addition, the set size was varied independently for the cued, relevant set of search items (e.g. red) and for the uncued, irrelevant set of search items (e.g. green). Based on the finding that observers can limit their search within the relevant set in a top-down manner (Donk & Theeuwes, 2001; Egeth, Virzi, & Garbart, 1984; Kaptein, Theeuwes, & van der Heijden, 1995), the subset-specific slopes and their difference were then used as a measure of selectivity. By using the postcuing method, and by measuring accuracy and the derived subset-specific slopes, the experiments showed that memory cues enhanced selectivity gradually up to 400 ms, after which it stabilized or declined again. These results thus provide strong evidence for that this is the time course of attentional selection as based on an internal target template.

The influence of the type of template on the time course of attentional selection

The aforementioned results were acquired by using visual cues picturing the target or target features. By such cues the studies have attempted to probe templates in visual working memory. A related question that has been repeatedly asked in these studies is whether attention can be allocated in a similar manner on the basis of categorical or verbal cues. This is not trivial as in the vast majority of everyday conditions attention needs to be directed in the absence of the target image, but based on a more categorical description, such as when searching for a pen of any kind on the table (Hollingworth, 2013; Schmidt & Zelinsky, 2009). The consistent finding from cued visual search studies has been that selection is always less efficient when based on a categorical cue than when based on a visual cue (Knapp & Abrams, 2012; Meyers & Rhoades, 1978; Schmidt & Zelinsky, 2011; Vickery, et al., 2005; Wolfe, et al., 2004). For example, in the study of Schmidt and Zelinsky, verbal descriptions of the target led to only a weak and slowly
increasing effect on the saccades that were directed to the target. While visual cues were found to improve saccadic accuracy rapidly, within about 400 ms, performance for the categorical cues did not change over time. These results suggest that visual working memory representations may be more effective, but they also raise the possibility that the superior effects of visual cues on search guidance may be at least partly due to visual priming, rather than an intentional working memory-based template (Vickery et al., 2005; Wilschut, Theeuwes, & Olivers, 2013b).

Chapter 6 compared the effects of target templates based on categorical and visual information, but controlled for cue-target priming by using neutral postcues. Again we measured search accuracies and subset-specific slopes. In the visual condition the precue consisted of the two color patches as described earlier, whereas in the categorical condition words were shown to indicate the two colors. A postcue pointed to one of the template colors (in the form of patches or words) that needed to be selected. The results showed that when we controlled for the effect of visual priming, the selectivity for the target color developed with an identical time course for both categorical and visual conditions. Also the overall accuracy measure (irrespective of set size) was indistinguishable for the two cue types.

In addition, Chapter 6 directly tested whether visual priming from cue to target would further improve search task performance. Now the neutral postcue was contrasted with a postcue which itself carried the actual target color. This colored postcue was found to increase overall accuracy relative to the neutral postcue, especially early in time, demonstrating that direct cue-target priming indeed also modifies performance in the cued visual search task. However, selectivity as such was unaffected by priming, as the subset-specific search slopes did not differ between the postcue types. In sum, the results thus show effects of priming on cued target search, and that when such priming is controlled for, visual and categorical cues lead to similar dynamics of selection - in contrast to earlier claims.

_is attentional selection based on internal templates sustained or transient?_

Similar to the spatial cueing effects as caused by external signals, the beneficiary effects of the template on selection may also be transient. Performance first increased but then decreased in the studies of Meyers and Rhoades (1978) and of Schmidt and Zelinsky (2011). Also the experiment of Vickery et al. (2005) showed that the cueing effect eventually diminished over time. Wolfe and colleagues (2004) suggested that a transient pattern was present in the RT slope measures in three of their experiments, even though
the decline could not be confirmed due to a lack of power. In the study of Knapp and Abrams (2012), a decrease in performance after 800 ms was reported for one of the three experiments. Notably also the Chapters 5 and 6 of the present thesis provide some support for the idea of transient attentional selection as based on internal target templates. Similar to the previous studies, the effect was relatively weak when each experiment was considered individually. However, if we collapse the results across experiments, selectivity proves to decline significantly, after the peak at 400 ms. Figure 2 shows the collapsed selectivity pattern, depicting the characteristic enhancement and the following decay (which was significant at \( p < .02 \)). It thus seems that, in contrast with the common conception of sustained biases, the time course of attentional selection by endogenous cues can be transient too.

![Image](image.png)

Figure 2. Mean search selectivity as a function of SOA, as collapsed across the three experiments of Chapters 5 and 6 that used visual precue and neutral postcue.

**Interim conclusions: Slower and transient attentional selection based on internal templates**

The results from studies employing cued visual search tasks are variable, as are the methods and response measures that have been applied. Chapters 5 and 6 of this thesis show how some of the previous results have likely been influenced by the type of measurement used (RT vs. accuracy) and by visual priming. The current experiments use a cueing method that controls for these confounds, and find that attentional selection develops within approximately half a second when based on an internal working memory
template. Furthermore, inspection of the existing studies reveals a rather consistent finding that such an attentional selection may not be sustained, as is typically postulated for endogenous attention, but declines over time. Thus, sustained endogenous attention may be possible when based on spatial information, while feature-based selection has a transient component.

**Neural explanations of transient cueing effects**

To account for the transience of attentional selection, it is helpful to consider the possible corresponding physiological mechanisms. In general, the transience of performance is reminiscent of the transient firing rate that is characteristic of the many neuronal populations when activated. For example, a rapid build-up that is followed by a decay is (even when the stimulus is sustained) characteristic of light-stimulated neurons in a number of visual processing areas (Schmolesky, et al., 1998). Similarly, higher order populations that are triggered by processing at earlier levels may initially respond strongly, after which either neural fatigue or regulatory inhibitory processes kick in. Such transient attentional responses may occur at cortical or subcortical levels.

**Cortical mechanisms**

The current thesis emphasizes the role of competition in shaping the time course of attention. Competition for selection has also been extensively described for target representations at a cortical level (Beck & Kastner, 2007; Kastner, et al., 2001; Reynolds, Chelazzi, & Desimone, 1999). When two or more stimuli are presented in the receptive field of a neuron, the response of that neuron is suppressed relative to when a single stimulus is presented. Such competition is automatic, occurs throughout the first feedforward sweep of visual processing, and depends on the size of the receptive field, its scope increasing with higher levels of processing with larger receptive fields (Kastner, et al., 2001). In addition, initial competition may take place via other inhibitory mechanisms, as suppressive interactions have also been found for items that fall in separate receptive fields (Allman, Miezis, & McGuinness, 1985; Lamme & Roelfsema, 2000). Thus, the competition in feedforward processing of the stimulus might contribute to the transient attention pattern. In Chapter 3 we found that overall task performance was impaired by distractors presented close to the targets, compared to when distractors were presented further apart, in line with the idea that target representation is more severely degraded if competition takes place already at the early processing areas where receptive fields are smaller. However, independent of distance (or other factors that increased competition),
the same transient component to performance was found. This suggests that competition at the earliest stages of processing as such does not explain the transient attention pattern.

Importantly, the transient attention pattern is time-locked to the cue, that should then trigger the modulation of the competed target representation. It has been found that competition occurring within a neuron’s receptive field can be biased in favor of the relevant input by attention (Desimone & Duncan, 1995; Moran & Desimone, 1985; Reynolds, et al., 1999). Also this biasing of competition can be found at different levels of processing, in different visual areas that vary in their neuron’s receptive field size (Beck & Kastner, 2009). Note that such biasing feedback signals can be triggered by both exogenous and endogenous attention. This feedback response may follow the transient bottom-up signal of the cue (as the bottom-up signal becomes stronger, the modulation it triggers also become stronger). Spatially specific cues may thus cause a rapid enhancement of target representation via relatively local feedback activation that occurs directly in visual processing areas (Lamme & Roelfsema, 2000).

In addition, feedback activation from higher processing areas may be inherently transient in nature. Selection based on internal top-down settings is presumed to involve circuits between the prefrontal cortex and more posterior areas that maintain the selection template (Desimone & Duncan, 1995; Hamker, 2004). When selection is based on location, the likely site of the template lies in parietal regions, where responses are found to be increased by spatial attention (Colby, Duhamel, & Goldberg, 1996; Greenberg, Esterman, Wilson, Serences, & Yantis, 2010; Serences & Yantis, 2007). In the case of attentional selection by target feature or object, temporal and occipital areas have been suggested to hold the template (Chelazzi, Miller, Duncan, & Desimone, 1993; Moro, Tolboom, Khayat, & Roelfsema, 2010; Motter, 1994). Neural modulations in these areas as a response to a search cue occur typically after 200 ms, which is not inconsistent with the behavioral effects. However, the activation related to template maintenance in prefrontal cortex or more posterior areas appears to be relatively sustained (Courtney, Ungerleider, Keil, & Haxby, 1997; Kelley, Serences, Giesbrecht, & Yantis, 2008; Wilson, O Scalaidhe, & Goldman-Rakic, 1993). The same goes for most endogenously cued performance. For example, as shown in Chapters 2 and 3, when no competing distractors are present, cueing effects are sustained. Therefore, while the frontoparietal or frontotemporal network may underlie the endogenous target selection, the transience is unlikely to stem directly from their activation pattern.
Instead, the data concerning the spatially specific cueing effects in this thesis (Chapters 2-4) are most consistent with a combination of two sources of feedback modulation: a transient local signal activation that enhances target representation directly in visual processing areas, in combination with a sustained feedback from higher-order areas. The transiently enhanced bottom-up signal activation provides the candidate for the higher level feedback to be directed at. While the initial bottom-up activation enhances the target signal in a spatially specific manner, it may not be fully able to filter out the simultaneous competition. Similarly, the top-down feedback as based on a spatial template is directed to the target, but it is then vulnerable to same competition. This feedback signal is sustained at an optimal level when there is only one object, i.e., the target. It fully takes over the initially bottom-up driven bias. However, in case of competition, some of this feedback, while still sustained, will inevitably leak to the other objects in the display, and thus becomes less effective. Therefore, by influencing the sustained feedback signal, competition allows the transient stimulus-based signal to be revealed. Figure 3 illustrates this scenario.

Figure 3. The figure depicts two possible combinations of a bottom-up transient signal and a top-down sustained signal as a response to a spatially specific cue, on the left without competition, and on the right with competition. Competition lowers the top-down feedback and the resulting activation pattern assumes a transient component.
INTRODUCTION

Enhancement originating from subcortical structures
In addition to the modulations at cortical level, target selection may be enhanced by input from subcortical sources, in response to cues. For instance, striatal dopamine has been suggested to play a role in the RSVP task in regulating the selection of items that will gain access to working memory (Colzato, Slagter, Spapé, & Hommel, 2008; Colzato, Slagter, de Rover, & Hommel, 2011; Slagter, et al., 2012). Similar to the RSVP task, dopamine could also modulate the target selection in cueing tasks.

It has also been suggested that the transient attention time course could be explained by phasic norepinephrine release from the locus coeruleus nucleus in the pons (Nieuwenhuis, et al., 2005). According to this locus coeruleus-norepinephrine (LC-NE) hypothesis, temporal properties of the phasic norepinephrine response underlie the attentional enhancement and subsequent decline as found in the RSVP task. The phasic LC-NE response is caused by task relevant items such as cues or targets, with a peak latency of about 100-150 ms (Aston-Jones & Cohen, 2005). After the rapid enhancement the response becomes refractory and cannot be activated for some few hundred milliseconds, corresponding with the target discrimination time course in RSVP. It could also more generally explain the transient attention time course as found in the cueing tasks.

Note that the modulatory effect of the phasic LC-NE response is in principle spatially non-specific, as NE is released throughout the cortex. Similarly, the effects of striatal dopamine appear widespread, as it has been assumed to regulate WM encoding via indirect projections to frontal cortex (Slagter, et al., 2012). Nevertheless, such non-specific modulations may lead to specific (i.e. spatially selective, or specific feature-based) effects if one assumes that the NE or dopamine response also enhances the current selection settings. A spatial cue may itself induce a specific local enhancement, which is then, by multiplication, further enhanced by subcortical modulations. Thus, rather than assuming a single neural correlate, selection is more likely subserved by the interplay between specific and non-specific mechanisms. This would also explain why a non-specific signal alone does not lead to transient selection effects (Chapter 2, Experiment 3), since there the signal is not selective. Further research is needed to resolve whether transient attention is produced by cortical or subcortical mechanisms, or their specific combination.
Chapter overview

The following chapters present studies that were conducted to investigate the time course of attentional selection as based on external signals (Chapters 2-4) and internal representations (Chapters 5-6).

Chapter 2 investigates how attention evolves over time when the selection focus remains constant at central fixation. The effects of selectivity are investigated by contrasting distractors present and absent conditions. The results show that attentional selection has a rapidly increasing and transient time course also when there is no spatial shifting involved. Importantly, for the first time they show that the transient pattern occurs only in the presence of competition, and when the target is presented alone, performance is sustained.

Chapter 3 further examines the conditions of competition, and generalizes the results of Chapter 2 to spatial cueing with location uncertainty. The findings indicate that competition underlies the transient attention time course also in these more typical conditions. The experiments further show that competition in general, at multiple levels of processing, results in the transient pattern. Overall performance is influenced by the strength of competition, but the time course depends solely on the presence of distractors.

Chapter 4 studies stimulus-based effects on the time course of selection. By manipulating the target and cue type in a spatially invariant task the experiments show that for clearly separable stimulus pairs the cue enhances target visibility with a rapid time course. However, when the target and cue are similar objects, their representations are prone to confusion and performance suffers at intermediate cueing times. The relative stimulus energy of the cue and the target has a further influence on the target visibility at very short cueing times. The combination of these early sensory and later attentional effects can make the performance time course appear either enhancing or declining.

Chapter 5 investigates the time course of attentional selection as based on internal target templates. Visual search target is cued by a neutral postcue that indicates the correct target in a memorized template. This procedure controls for cue-target priming and allows for examining selection as a function of mere WM template. RT task reveals no modulations in selectivity, while an accuracy task shows a gradually building enhancement that peaks at 400 ms. This is suggested to reflect the time course of selection as based on memory templates.
Chapter 6 uses the method of Chapter 5 to test the influence of template type on the time course of attentional selection. A direct comparison of direct visual cues and memory cues show that both lead to a similarly efficient search guidance. However, overall accuracy performance is enhanced by cue-target priming, and this effect lasts at least for the first half a second. In addition, categorical and visual template cues are found to result in identical search performance. These results then challenge the earlier conception that visual search is guided efficiently only when based on visual WM templates, by showing that some of the effects can be accounted by visual priming and methodological problems.