Doctoral thesis submitted to
the Faculty of Behavioural and Cultural Studies
Heidelberg University
in partial fulfillment of the requirements of the degree of
Doctor of Philosophy (Dr. phil.)
in Psychology

Title of the thesis

Goal-Directed Visual Search: The Role of Cognition, Motivation and Emotion

presented by
Tanja Kollei

year of submission
2014

Dean: Prof. Dr. Klaus Fiedler
Advisor: Prof. Dr. Andreas Voss

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

I am indebted to many people who have contributed to this thesis by giving me their support, advice and encouragement:

First and foremost, I would like to thank my advisor Prof. Dr. Andreas Voss for his invaluable help, advice and guidance whenever I needed it. I am especially grateful for the leeway that he gave me in conceptualizing the research questions and for his encouragement to pursue them despite some challenges.

I also express my gratitude to Prof. Dr. Ursula Christmann who gave me her advice on many scientific matters, who encouraged me and who kindly agreed to review my thesis.

I thank all the participants that took part in the experiments and Miriam Hans, Ina Knittel, Ulf Mertens, Jasmin Munske, Sophie Schoenmakers, Michael Spektor and Imme Zillekens for data collection. I especially thank Miriam Hans for the efforts that she put into the recruiting of participants and for testing all of the experiments.

My colleagues Markus Nagler, Andreas Neubauer and Veronika Lerche enriched this thesis with many helpful comments on the conducted experiments – thanks to all of you! I also thank Markus Nagler for testing some eye tracking experiments and for his support during our “C lessons”.

I thank my colleague and “room mate” Veronika Lerche for patiently sitting in front of the eye tracker and testing some of the experiments, for proof-reading earlier versions of this thesis and for the many inspiring and entertaining discussions – I hope we can soon start investigating our many new research questions!

Special thanks go to Martin Jost for proof-reading and to Carla Minarik for proof-reading and for helpful comments and enlightening talks during our many “S-Bahn” rides. Thanks also to Rosalux Falquez for her advice and support.

I am deeply grateful to my sister Ines Kollei who always supported and encouraged me and who gave me valuable feedback on earlier versions of this thesis.

Finally, this thesis would not have been possible without the support of my family throughout the last years and before. Above all, I thank Mohammed for always reminding me of what matters most.

1 INTRODUCTION

Since the 1950s, cognitive psychology has defined attention as the central bottleneck through which information has to pass before it can be further processed and acted upon. Our senses process thousands of stimuli simultaneously, but what we perceive is always only a selection of those stimuli that our attention system classifies as the most relevant for the current situation (Yantis, 2000). Attention is the precondition for almost any kind of cognitive operation or behavior, which is why it is an interesting subject for various disciplines of psychology. Whereas attention research in the 1950s focused on auditory attention, current research predominantly studies visual attention because visual stimuli can be manipulated more conveniently and visual attention can be measured more reliably (Pashler, 1998). The visual search task has become the most influential method to measure visual attention.

Due to the rise of cognitive psychology, many cognitive theories of depression and anxiety conceptualize an attention bias as an important cause and/or effect of these dysfunctions (Yiend, 2010) and a growing body of research studies how attentional trainings can help to alleviate symptoms of depression and anxiety (Wadlinger & Isaacowitz, 2011). There is abundant evidence which suggests that visual attention is heavily influenced by participants' motivational and emotional state and that attention plays a central role for the self-regulation of behavior. Despite the fact that almost any article on visual search acknowledges this by beginning with an example of visual search in the real life [e.g., "When driving along a city road, the sudden appearance of a child from behind a parked car may grab our attention." (Belopolsky & Theeuwes, 2010, p. 2543)], theories of goal-directed visual search still explain search and goals with purely cognitive terms (e.g., Leber & Egeth, 2006; Theeuwes, 2010; Wolfe, 1994; Zehetleitner, Goschy, & Müller, 2012). This is probably motivated by their aim to describe attentional selection as a general concept and to arrive at conclusions with high internal validity. Motivational and emotional aspects are implicitly considered confounding variables that obscure the "true" cause of the effect. This doctoral thesis argues that the sole focus on cognitive processes of attention is problematic for theories of goal-directed search. By not considering the motivational process of goal striving, these theories only study a fragmentary concept of voluntary attentional control in the real life. Thus, inferences about general processes are not possible. Theories of visual search have to consider the factors that determine the strength with which a goal is pursued in order to find out whether and how goals can guide attention.

Previous research has focused on unintentional influences of motivation on visual search. Based on evolutionary psychology, the *threat superiority hypothesis* assumes that attention is involuntarily captured by motivationally meaningful stimuli (e.g., threatening faces) and that this process cannot be controlled by task goals (e.g., Öhman, Lundqvist, & Esteves, 2001). This proposition contradicts most theories on visual search, which argue that only a limited set of features can guide attention without being mediated by goals. Furthermore, the empirical evidence supporting the threat superiority hypothesis is weak because most studies did not consider the fact that search for emotional stimuli was always relevant for the task goal. Accordingly, this thesis aims to demonstrate that stimuli with motivational content only affect visual search if this content is relevant for the observer's goal.

To conclude, this thesis argues that both the motivational/emotional and the cognitive perspective on attentional control in visual search can benefit from each other. In order to study the preconditions of goal-directed guidance of attention, it is important to consider the factors that determine the strength with which a goal is pursued. In the same vein, the experimental design has to ensure that any emotional effect is not mediated by task goals in order to examine whether emotional stimuli are automatically processed during visual search.

Chapter 2 outlines the theoretical background of this thesis by providing an overview of some leading theories of goal-directed control in visual search (see Chapter 2.1.) and by showing that cognitive research on selective attention can profit from motivational accounts of goal-directed action. The central argument is that goal-directed control in visual search depends, at least partly, on the motivation of the observer so that a seemingly cognitive deficit in visual search might actually be a motivational one (see Chapter 2.2). On the other hand, it is also discussed how a seemingly motivational effect (attentional capture by emotional faces) could indeed be a cognitive one (see Chapter 2.3). **Chapter 3** describes the aims of four experimental studies, derived from the theoretical background. **Chapters 4 - 7** present the hypotheses, methods and results of those studies. Each chapter closes with a discussion of the results. **Chapter 8** provides a comprehensive summary and discussion of the results of the four experiments.

2 THEORETICAL BACKGROUND

2.1 Goal-Directed Visual Search: the Cognitive Perspective

2.1.1 A Short Introduction to Selective Attention and Visual Search

The selection of relevant stimulus input is the core function of attention and it is the core question in attention research how this selection process is controlled. Donald Broadbent's seminal *filter theory* (1958) introduced a basic model of attentional selection to cognitive psychology that has been adopted and modified by all major theories of attentional selection since. His theory conceptualized perception as a two-stage process: In the first *preattentive* stage, physical properties of all of the incoming sensory signals are extracted in parallel. Processing during this stage is not limited by cognitive resources. In the second, *postattentive* stage, only a selection of those stimuli is processed more elaborately. Physical characteristics that have passed the filter are transformed into cognitive, meaningful constructs and are perceived. Attention is the process of selection from the first to the second stage and is necessary because the limited capacity of the second stage would otherwise suffer from cognitive overload (see also Driver, 2001). Each theory on attentional selection in visual search that is discussed in the following chapters incurred this basic distinction between a parallel, preattentive and a more elaborate, postattentive, stage of processing.

Allport (1989) delivered an interesting view on the function of the bottleneck of attention by emphasizing the link between selective attention and action control. He argued that the limited information-processing capacity of the human brain is not the actual cause for selective attention to occur. Instead, attention has to be selective because perceiving every sensory signal evoked from the surroundings would preclude adaptive, goal-directed behavior. To initiate and execute an action, only those objects in the visual field that are relevant for this particular action need to be processed while irrelevant information that could interfere with the execution has to be inhibited. For example, when an individual is trying to catch a ball that is flying in the air, attention networks have to selectively transmit the position, distance and size of the ball in order to coordinate the movement. Other objects in the visual field might also have to be processed in a limited way in order to gather information, but attention has to ensure that they do not elicit an action at the same time – if it is not necessary. Based on this reasoning, Allport termed attention's core function as *selection-for-action*. As many stimuli in the environment might require a reaction or pose a chance to achieve the current goal, attentional selection is a precondition for action control. At a certain point in time, only a limited amount of information can be processed because the organism can only execute a limited amount of actions simultaneously. Thus, attentional control is an essential component of goal-directed behavior and self-regulation (Bodenhausen & Hugenberg, 2009).

All theories of attentional control converge on the view that it consists of the interplay of *stimulus-driven* (also termed *bottom-up*, or *exogenous)* and *goal-directed* (also termed *top-down* or *endogenous*) processing. Goal-directed selection is driven by the *intentions* of the observer and bottom-up selection occurs whenever participants' attention is *unintentionally* directed to stimuli. Stimuli evoke high bottom-up activation in the visual system if their physical characteristics (color, motion, orientation, and size) are very different from the surrounding stimuli (e.g., a person wearing a white shirt in a crowd of people wearing blue shirts) (Corbetta & Shulman, 2002; Desimone & Duncan, 1995; Egeth & Yantis, 1997; Theeuwes, 2010; Wolfe, 1994; Wolfe & Horowitz, 2004; Yantis, 2000). Both types of control are preconditions for effective action control. Bottom-up activation is important because it enables individuals to adjust their current action quickly to changes in the environment. While focusing on the moving ball in the air in order to catch it (top-down goal), distracting, irrelevant stimuli must be ignored (e.g., a bird flying high in the sky), whereas distracting, but relevant information must be detected (e.g., another approaching ball that might hit one). Thus, goal-directed action requires a complex interplay of stimulus-driven and goal-directed attentional control and attention research aims to explain how this interplay comes about. Current theories on attentional control diverge in their propositions about this interaction and each of them is supported by a lot of empirical studies. Therefore, attentional control is a highly debated issue and this theoretical background introduces three theories that contributed much to the debate: One theory argues that bottom-up and top-down control processes operate in strictly separate stages of visual selection. Other theories, however, maintain that both control processes work together in all of the stages of visual selection. The following description of visual search illustrates what the different stages of selection refer to.

One of the most important methods to study attentional control is the visual search task (for an overview of other methods see Fox, Derakshan, & Standage, 2011). Participants see several items arranged on a screen and have to find and identify a target. This task allows studying if attentional deployment is guided by the goal of finding the target, by the irrelevant physical characteristics of the stimuli in the display, or both. The theories of visual search that are discussed in the following chapters are based on the most influential theory of visual search, *feature integration theory* (Treisman & Gelade, 1980) and adopted its central concepts. *Features* are the basic physical attributes of objects; they are the building blocks of perception. The visual system groups these features into *dimensions,* such as color, motion, orientation, and size (Wolfe

& Horowitz, 2004). For example, blue is one of many features in the dimension color. All of the theories assume that each feature is processed in a different channel. The bottom-up activation of a stimulus is computed by calculating the difference between the feature characteristics of the stimulus and those of the surrounding stimuli. The greater this difference, the greater the bottom-up activation of the stimulus. A *salient* stimulus is a stimulus that causes higher bottom-up activation in the visual system than its surroundings. This is the reason why a person wearing a white shirt is easy to detect in a group of people wearing blue shirts.

Feature integration theory (Treisman & Gelade, 1980) adopted the distinction between two separate stages of processing from filter theory (Broadbent, 1958). In the first, preattentive stage, the whole visual field is scanned in *parallel*. Information about the strength of a feature signal at a certain location in the visual field is saved in distinct *feature maps.* Targets that are defined by one feature (e.g., color) can be detected and identified in this stage. For example, a red “T” among green “T”s is found very fast because its defining feature green pops out from its background. As this process does not require the focusing of attention, it is called preattentive. Search is not restricted by capacity limits and all items in the visual field are processed at the same time. The second stage of *serial* focal attention is necessary if the target is defined by a conjunction of features, for example a green “T” among brown “T”s and green “X”s. Neither the orientation (“T”) nor the color (green) of the conjunction target pops out because these features are shared by some distractors. The identification of such a target requires the binding of separate features by focusing attention on each item serially. Therefore, processing during this stage needs more time. The transfer of information from the first stage to the second stage is the bottleneck where visual selection takes place. Based on the feature maps computed in the first stage, focal attention is directed towards the location with the highest feature activation in the second stage and identifies it. After a phase of parallel, unlimited information processing, only one item at a time can be processed in the serial stage of search. Accordingly, response latency for the target is independent of the number of stimuli in preattentive search tasks, whereas it increases linearly with each additional stimulus in tasks that require feature binding. Thus, the former are also called parallel and the latter serial search tasks.

Many leading accounts of visual search adopted the distinction between a first stage of parallel or preattentive search and a second stage of serial or focal search (Driver, 2001). The classic method to differentiate between these two stages is the computation of search slopes, which requires the manipulation of the number of items in the display (e.g., one condition contains five items and the other condition contains eight items). The increase in mean response latency from the condition with fewer stimuli to the condition with more stimuli is divided by the additional number of stimuli (in this example: three). The resulting search slope indicates the additional search time caused by each additional item in the display. It is common to categorize search as parallel if search slopes are smaller than 10 ms/item and to categorize search as serial if search slopes are larger than 10 ms/item (Müller & Krummenacher, 2006b). This dichotomy plays a central role in the theories which are discussed in the following chapters (see 2.1.3 - 2.1.5). Whereas one theory argues that the first stage of parallel processing can only be guided by bottom-up information, other theories propose that top-down information modulates both stages of processing. Therefore, the question of whether participants' goals can guide search during the parallel stage is up for debate. Before the propositions of these theories are presented, the next chapter (see 2.1.2) discusses the measurement of visual search in experimental psychology.

2.1.2 The Measurement of Visual Search: Covert and Overt Attention

Visual search is traditionally observed by measuring response latency, that is, the time between the onset of the stimulus display and the manual response that is given by the observer once she has detected the target. The faster the manual response, the faster the target is found among the distractors. Since the development of easy-to-use and affordable eye tracking technology in the 1980s and 1990s, eye movements during search can be analyzed as dependent variables, as well. This led to the differentiation between covert and overt attention: Covert attention refers to visual selection that does not involve eye movements whereas overt attention means that visual selection requires eye movements (Findlay & Gilchrist, 2003). Experimental research typically differentiates between saccades and fixations: The former are fast movements of the eyes from one location to another and the latter describe moments in time when the eye is fixated on a specific location and does not move.

Response latencies and accuracy are only global measures of visual search and do not allow making inferences about the sub-components of search. Furthermore, one can only determine whether search is parallel or serial search with these dependent variables if set size is manipulated. If response latencies are positively correlated with the number of items in the display, search must be serial, whereas a zero correlation is characteristic of parallel search. However, serial search also consists of a first phase of parallel search (Treisman & Gelade, 1980). If a distraction effect during serial search is observed in response latencies, one does not know if the distraction occurred during the preattentive or during the serial stage. The monitoring of eye movements, however, delivers fine-grained information about the spatial (which locations in the display were fixated?) and temporal (when and for how long was a stimulus fixated?) characteristics of search and allows dissecting the preattentive and serial stage of visual search without varying set size. Beside these obvious advantages of analyzing eye movements, eye movements are simply necessary for many search tasks.

Eye movements are important for visual perception because visual acuity is restricted to a very limited area of the visual field. It is highest in the foveal region, which extends to an angle of approximately 1°. The parafoveal region spans from 1° to 5° and everything that is farther away belongs to the peripheral region (Findlay & Gilchrist, 2003). In daylight, visual acuity drops linearly from the foveal region to the border of parafoveal vision (5°) and from there it declines sharply, with an approximate exponential function. Visual acuity at 5° is only roughly 50% of the acuity in the foveal region (Duchowski, 2007). Eye movements are necessary if several items are dispersed in the visual field and if foveal fixation is needed to identify them.

According to the *active vision* account by Findlay and Gilchrist (2003), eye movements are an integral part of attention:

> We have pointed out that much thinking in passive vision implicitly downplays the role of the fovea. We make the counterargument that the radial organisation of the visual system based on the fovea is far from co-incidental but is rather its most fundamental feature. A simple but telling argument considers a hypothetical brain, which provided the same high resolution as found in human foveal vision at all locations in the visual field. It has been calculated that such a hypothetical brain would be some hundreds of thousands times larger than our current brain and so would weigh perhaps ten tons. A mobile eye constructed on the principles of the vertebrate eye is not a co-incidence or a luxury but is very probably the only way in which a visual system can combine high resolution with the ability to monitor the whole visual field. (Findlay & Gilchrist, 2003, p. 5)

The authors criticize the neglect of eye movements by *passive vision* theories. With this term, the authors refer to many theories of visual attention and visual search, amongst them feature integration theory. Findlay and Gilchrist argue that feature integration theory does not consider the sharp decline of visual acuity in the periphery of the visual field as a relevant aspect of the search process. The empirical studies on which this theory is based disregarded this fact by only using tasks in which the target could be detected without making eye movements. Therefore, stimuli are positioned very close to the fixation center in the typical visual search task. The authors maintain that the data generated with such tasks does not permit drawing conclusions about visual search because they do not measure visual search if eye movements are not involved. In the huge majority of search tasks that are required to prepare actions, the eyes have to move over the visual field to analyze and identify the elements in the periphery. Consequently, each theory on visual search should be based on observations of overt attention in order to generate valid propositions.

One obvious reason for the focus on covert attention was certainly that the measurement of eye movements was very expensive, time-consuming and unreliable until new eye tracking technology had been invented in the 1980s. Findlay and Gilchrist (2003) name another reason by arguing that cognitive psychology has been, especially in its early years, a science that aimed at studying only mental processes. Accordingly, many cognitive psychologists had no interest in eye movements because they regarded them as actions and therefore not as a relevant topic of cognitive research.

The active vision account (Findlay & Gilchrist, 2003) proposes that the function of covert attention is to prepare eye movements: "*The fixation act is the process of paying attention* and is supported by covert processes that result in peripheral preview for the next fixation location." (p. 40). Like feature integration theory, the account proposes that covert attention spreads in parallel across the visual display to compute saliency signals which are saved in an internal saliency map. A saccade is initiated to the location with the highest activation on this map. Thus, saccades are, of course, executed serially, but are programmed in parallel. Only one saccade to a specific location can be executed at any time, yet before the execution of this saccade has been finished, the next saccade is already being programmed. Based on this reasoning, the integration of eye movements into theories of visual search is quite easy because factors that are assumed to determine the control of covert attention should also affect overt attention.

The active vision account is supported by studies showing that the center of a fixation indicates the focus of attention (Deubel & Schneider, 1996; Hoffman & Subramaniam, 1995; Rizzolatti, Riggio, Dascola, & Umiltá, 1987; Shepherd, Findlay, & Hockey, 1986) and by theories and studies which argue and demonstrate that saccades are programmed based on a saliency map which is computed by parallel processing of the display (Baldauf & Deubel, 2008; Findlay & Walker, 1999; Godijn & Theeuwes, 2002; Trappenberg, Dorris, Munoz, & Klein, 2001).

Eye Tracking Technology and the Classification of Eye Movements

Zelinsky (2008) argued that the obvious advantage of the analysis of overt attention is that "[e]ye movements are directly observable" (p. 788). On the one hand, this is true because the camera of an eye tracker can indeed record the movement of the eye directly. On the other hand, the recorded center of the pupil first has to be mapped to coordinates on the computer screen and then eye movements have to be detected based on the multitude of gaze coordinates before the data can be analyzed quantitatively. Both processes require complex algorithms which often work reliably, but whose results of course can be distorted by measurement errors. These can originate in behavior of the participant (head movement or blinks) or in errors in the algorithms themselves. Therefore, the characteristics of the eye tracking technology play an important role in the measurement of overt attention.

All of the leading eye tracking manufacturers (Tobii Technology, SensoMotoric Instruments, SR Research) employ the video-based pupil-and-corneal-reflection method to record eye movements. This has become the most widely used technique since the 1990s (Holmqvist et al., 2011). Eye trackers that operate with this technology basically consist of a highly specialized video camera and an image analysis software. The eye tracker emits infrared light and its built-in video camera records the reflection of this light. Typically, the pupil appears dark on the recorded image because it absorbs the infrared light, whereas the other parts of the eye reflect it. Infrared light is reflected strongly from a location on the cornea, resulting in the corneal reflection. This reflection is important in order to distinguish small head movements from eye movements. The video of the eye is presented on a computer in real time. The image analysis software detects the center of the pupil and the corneal reflection in the recorded image of the eye, based on their visual characteristics. The locations of the eye that are classified as being the pupil and the corneal reflection are marked with crosses in the video, so that the experimenter can check if the software has correctly identified the pupil. In case of a correct detection of the pupil and the corneal reflection, the eye tracker can measure eye movements. In order to map these movements to coordinates on the screen (and thereby to locations on the display), a calibration procedure has to be conducted prior the beginning of the measurement. During the calibration, typically five, nine or 13 points are presented one after another in the center and in the corners of the screen and have to be fixated by the participant (Duchowski, 2007; Holmqvist et al., 2011).

Once the calibration procedure has been completed, the eye tracker computes the gaze position from the recorded eye video. Each recorded picture of the eye with the corresponding gaze position is a data sample. The total number of data samples is determined by the temporal resolution, measured in Hertz. An eye tracker with a sampling rate of 1250 Hz (currently the highest temporal resolution available by SensoMotoric Instruments and most other manufacturers and the one used for the experiments in this study) records gaze position 1250 times in one second. This results in approximately 1000 raw data samples during a typical visual search trial. Each raw data sample only contains information about the position of the fovea on the screen, but it does not reveal whether this data sample is part of a fixation or of a saccade. The application of a special algorithm is necessary in order to detect these eye movements in the raw data. The algorithms included in the software by the leading eye tracking manufacturers share two characteristics: First, they classify eye movements in *saccades* (the fast movement of the eye from one position in space to another), *fixations* (fixating the fovea for a certain point of time on one specific region in space) and *blinks* (periods of time when gaze is not recorded because the pupil is covered). Second, these algorithms use saccadic velocity and saccadic acceleration as the key criteria for event detection, given that the sampling rate is high (> 200 Hz). This means that the first step of analysis detects saccades in the raw data and that the second step categorizes the remaining raw data samples as being either part of fixations or part of blinks (Holmqvist et al., 2011).

Saccades are rapid movements of the eyes (and the fastest movements that the human body can execute) that typically last 30 - 80 ms (Holmqvist et al., 2011). People make 3 - 4 saccades per minute in most visual tasks. During the execution of a saccade, information uptake is not possible (*saccadic suppression*) (Findlay & Gilchrist, 2003). Eye movements that are faster than 30 - 100°/s and exceed an acceleration criterion of 4000 - 8000°/s^2 are typically classified as saccades by most event detection algorithms. The comparability of eye tracking studies suffers from the fact that the eye tracking manufacturers employ different algorithms for event detection (e.g., the event detection algorithm by SensoMotoric Instruments uses only saccadic velocity to detect saccades, whereas the algorithm by SR Research uses both saccadic velocity and acceleration) and that researchers can change some settings by themselves, which further increases the variability among event detection methods (Holmqvist et al., 2011). The latter poses a problem because there is no commonly accepted rule for the setting of these thresholds, so that even studies using the same software can vary in the pattern of eye movement data they compute.

Fixation durations typically vary between 50 and 250 ms in most experimental settings, but can also extend to several seconds, depending on the task. Fixation dispersion ranges from 0.5° to 2.0°. The saccade detection algorithm by SensoMotoric Instruments allows for the additional setting of a fixation duration threshold. A reasonable cut-off value is 50 ms, so that only those fixations during which visual processing takes place are included in the data sample. Fixation durations below 50 ms often represent eye movements that are falsely detected as fixations due to glissades (Holmqvist et al., 2011).

The classification of gaze data into saccades, fixations and blinks is sufficient for the vast majority of visual search studies because saccades and fixations are those eye movements which are of theoretical value for visual search and with which dozens of other variables can be computed. Yet, this binary classification of eye movements disregards, at least in some cases, that other eye movements exist as well. Fixations are indeed movements like saccades because the eye does not rest completely still during a fixation, but moves at a much lower speed as during a saccade. Therefore, a fixation also consists of so-called *microsaccades,* whose function it is to stabilize the eye over the center of fixation once it has drifted away. *Glissades* are small eye movements around the fixation location after the end of the saccade and are normally treated as noise in the data because they complicate the detection of saccades and fixations. *Smooth pursuit* is a continuous movement of the eye that follows a specific object (e.g., following a soccer player running along the field). This movement is much slower than a saccade and also driven by a different brain region than saccades (Holmqvist et al., 2011). Therefore, it is of little interest to research on visual search, for which the coupling between covert attention and the execution of saccades is of main interest. The differentiation between microsaccades and fixations and between glissades and saccades or fixations is not implemented in the standard event detection algorithms delivered by the manufacturers. The existence of microsaccades does not pose a problem for the detection of fixations because the latter are defined, in high-speed event detection, by the absence of a saccade. As microsaccades are much slower than a saccade, they fall below the saccadic detection threshold. Glissades can be problematic because they erroneously lead to the detection of very short, “false” fixations (1 – 10 ms). This problem can be alleviated by setting the minimum fixation duration to 50 ms.

Key Eye Movement Variables in Visual Search

Once one has let the software detect fixations and saccades, one can compute dozens of dependent variables based on these eye movements [for an overview of possible and interesting variables see Holmqvist et al. (2011)]. These variables allow a fine-grained analysis of the visual search process and a distinction between the parallel and the serial phase of visual search. This chapter describes only those variables that are used in the experimental studies conducted for this doctoral thesis:

- *number of fixations executed during a trial*
- *hit rate of the first saccade on the target/distractor*
- *saccadic latency*
- *fixation duration*
- *hit rate of all saccades on the target / distractor*

Number of Fixations executed during a Trial

The number of fixations that participants execute on average during a trial is an indicator of the efficiency of visual search. D. E. Williams, Reingold, Moscovitch, and Behrmann (1997) investigated eye movement patterns during parallel and serial search. In the parallel search task, participants had to search for the letter “O” among “X”s and in the serial search task for the letter “T” among “L”s. Thus, the former was a classic feature search task, in which the target was supposed to be found with parallel search, whereas the latter was assumed to require the conjunction of features and the serial allocation of attention. The results confirmed this hypothesis concerning manual reaction time: Search slopes were flat for the feature search task and increased linearly with an increase in set size for the conjunction task. Importantly, participants constantly needed on average two fixations in each trial of the parallel search task. In the serial search task, however, the number of fixations increased linearly, like the response time. Similar results were reported by Zelinsky and Sheinberg (1997) and Scialfa and Joffe (1998). The number of two saccades seems to be a reasonable cut-off to differentiate between parallel and serial search because it is consistent with the active vision account and related theories (Baldauf & Deubel, 2008; Findlay & Gilchrist, 2003; Findlay & Walker, 1999; Godijn &

Theeuwes, 2002; Trappenberg et al., 2001), which assume that two saccades can be programmed in parallel.

Therefore, eye tracking allows making inferences about the stage of the search process (parallel vs. serial) without varying set size and computing search slopes. However, the number of fixations is not used as an indicator of parallel/serial search in any of the studies that are cited in the next chapters. Even if eye movements are monitored, most experiments do not report the average number of fixations needed to complete search and instead focus on the first saccade (because this is the variable that should be most sensitive to processes during parallel search). One reason for this is surely that the analysis of search slopes has always been done before and is regarded as the state of the art method. Consequently, there has not developed any research tradition in interpreting the number of saccades as an indicator of parallel or serial search.

Hit Rate of the First Saccade on the Target / on the Distractor

The location of the first saccade is an indicator of the initial focus of attention and therefore the most important dependent variable in many visual search studies. After the parallel scanning of the whole visual display, the first saccade is directed towards the location with the highest bottom-up and/or top-down activation (Findlay & Gilchrist, 2003; Godijn & Theeuwes, 2002). Therefore, the percentage of trials in which the first saccade hits a certain object (e.g., target or distractor) is an indicator of the saliency of this object. If the hit rate of the first saccade on the target is significantly higher than that on the distractors, the target is more salient than the distractors during the preattentive phase. If the hit rate on the distractor is equal to that on the target, both exert a similar influence on attentional control. The hit rate of the first saccade on the target also reveals how efficient the search process is: The higher the hit rate on the target, the easier it is for participants to complete search in the parallel stage. Parallel search is accompanied by a relatively high percentage of first saccades towards the target, whereas a low percentage is characteristic for serial search. Unlike for search slopes and the number of fixations, there exists no cut-off with which to differentiate parallel and serial search based on this variable.

Saccadic Latency

Saccadic latency is the time between the onset of the stimuli and the initiation of the first saccade on the target. It is the temporal counterpart of the hit rate of the first saccade. Whereas the latter indicates the location in the display which has received the highest activation on the saliency map during parallel processing, saccadic latency indicates how long the processing of the saliency signals takes. The faster a correct first target saccade is initiated, the faster parallel attention has detected the target correctly. Saccadic latencies typically range from 100 ms to 300 ms, depending on the task (Holmqvist et al., 2011). Saccadic latencies that are shorter than 80 ms are categorized as anticipatory saccades. These are normally excluded from data analysis because they must have been initiated before target onset.

Fixation Duration

Based on the study by Just and Carpenter (1980) on eye movements during reading, fixation duration is used as an indicator of the depth of visual processing: The longer the duration of a fixation, the more elaborate the processing of the information at the location of the fixation. Holmqvist et al. (2011) argued that this interpretation can lead to false conclusions because fixation duration is an indicator of two processes: the mental processing of the stimulus at the fixation location (discrimination task) and the preparation of the next saccade (selection task). Yet, Findlay and Gilchrist (2003) proposed, based on findings reported by Hooge and Erkelens (1999), that fixation duration is mainly determined by the discrimination task. This is in line with the active vision account and related theories (Baldauf & Deubel, 2008; Findlay & Gilchrist, 2003; Findlay & Walker, 1999; Godijn & Theeuwes, 2002; Trappenberg et al., 2001), which argue that saccades are programmed in parallel. The programming of the saccade may still take place during the fixation, but it has begun way before. Fixation duration is an important variable to investigate whether a delay in response latency in one condition (e.g., in which a salient distractor appears in the display) is due to the *shifting* of attention (i.e., a saccade often falls on the distractor) or if it is due to a failure to *disengage* attention from one location to another (i.e., once a saccade is made to the distractor, the distractor is fixated longer than other items). This differentiation is central to a very influential hypothesis in research on attention bias towards emotional stimuli. According to the *delayed disengagement hypothesis* (Fox, Russo, Bowles, & Dutton, 2001) threatening stimuli do not affect the shifting of attention to a greater extent than neutral stimuli, but prolong the disengagement of attention. Fox et al. (2001) measured delayed disengagement indirectly via response latencies in a modified spatial cueing paradigm (see also Fox, Russo, & Dutton, 2002). Recently, Belopolsky, Devue, and Theeuwes (2011) accumulated evidence for this hypothesis by measuring delayed disengagement directly via the fixation duration on the threatening stimulus.

Hit Rate of All Saccades

Whereas the hit rate of the first saccade delivers specific information about the location that has received the highest activation during parallel processing of the display, the hit rate of all saccades is a global indicator of attentional selection: It is the percentage of trials in which any of the saccades that were executed during a trial (i.e., the first, the second, the third or a later saccade) falls on a certain stimulus. Thus, it shows whether and how often a certain stimulus has been hit by an observer's gaze sometime during the trial. More important, it also reveals participants' performance in finding the target correctly: If the hit rate of the first saccade on the target and the hit rate of the first saccade on the distractor do not differ, this suggests that both items were equally salient during the parallel stage of processing. If the hit rate of all saccades on the target is increased in comparison to the hit rate of the first saccade on the target, this shows that participants increased their search performance during serial search. In the same vein, if the hit rate of all saccades on the distractor is not significantly increased in comparison to the hit rate of the first saccade on the distractor, the distractor only captured overt attention during the first phase of parallel orienting, but not later on.

Therefore, eye tracking allows measuring the different components of attention: the first phase of parallel search, the depth with which different items are processed and the serial deployment of attention. This is clearly an advantage over response latency and accuracy, which are only global indicators of visual search. If the analysis of search slopes and/or of the number of fixations demonstrates that search was serial in a task, the hit rate on the target/on the distractors and saccadic latency still deliver information about the parallel component of search.

The following chapters present the main propositions on the influence of goal-directed attentional control during parallel search by three leading theories. These theories were chosen as most relevant for this doctoral thesis because they accumulated empirical evidence in favor of their assumptions and against the arguments of competing theories with the same visual search task. Theeuwes (1991, 1992) first used the *additional singleton task* to test his *stimulus-driven capture account*. According to this account, parallel visual processing is solely affected by bottom-up activation and top-down control is not possible in this stage (see Chapter 2.1.3). This claim is countered by two theories arguing for top-down modulability during parallel search: the *search mode account* (see Chapter 2.1.4) and the *dimension weighting account* (see Chapter 2.1.5).

Another leading theory - the *contingent involuntary capture theory* by Folk, Remington, and Johnston (1992) - also maintains that goal-directed control during parallel processing is possible. This theory was deliberately not included in the theoretical background for this doctoral thesis because it was developed and tested with the modified spatial cueing task. This task does not allow measuring eye movements because the display is presented for only 100 to 200 ms (see also Theeuwes (2010), for an explanation of further procedural differences between the additional singleton task and the spatial cueing task). The theories that are discussed in the following chapters propose several determining factors of top-down control and it is difficult, if not impossible, to control all of these factors in one single experiment. Comparing these theories to a theory that was tested with a very different task would require controlling for additional factors, which would go beyond the scope of this thesis.

The theories of visual search that are discussed in the next chapters were mainly tested with tasks that only measured covert attention. All of them implicitly or explicitly build on Treisman's (1980) feature integration theory that conceptualized search as a covert process. Furthermore, all of them were developed in the 1980s and 1990s, when eye tracking was possible, but not with such an ease as nowadays. Consequently, they consider goal-directed visual search as a mainly covert process and the critique of Findlay and Gilchrist (2003), which was levelled at feature integration theory and at passive vision accounts, refers to them, too. Two of the accounts (see Chapter 2.1.3 and 2.1.5), however, have also been tested with eye movement studies.

2.1.3 The Stimulus-Driven Capture Account

Jan Theeuwes (for a review of his theory see: Theeuwes, 2010) developed a very influential theory that maintains that stimulus-driven activation is the determining factor of visual selection in the preattentive stage of visual search. Like feature integration theory, his theory proposes that local feature contrasts are calculated in parallel during the preattentive stage and that the item with the highest feature contrast (i.e., the most salient item) is selected for response and passed on to the second stage of perceptual processing. The central and controversial claim is that this process cannot be modulated by voluntary control and depends solely on the physical characteristics of the stimuli. Therefore, Theeuwes termed the first selection of the most salient item in the preattentive stage as *attentional capture*, thereby underlining the lack of intentionality during this stage: "This automatic shift of attention which is not the result of any top-down set on part of the observer is what is known as 'attentional capture'" (Theeuwes, 2010, p. 79). It is important to note that attentional capture is, by definition, always automatic in the sense that it occurs unintentionally. The item which captured attention is identified by top-down knowledge in the second stage and a decision is made whether it is the item one was looking for or not. If the most salient item is not the target, top-down control causes a quick disengagement from this item and a reorientation of attention.

Strong evidence in support for the stimulus-driven capture account was accumulated with the additional singleton task. In one of the most commonly used versions of this task (Theeuwes, 1992), the display consists of a shape target (e.g., a circle) among several shape distractors (e.g., diamonds). The target is a singleton in the display because it is distinct in its orientation and form from the distractors. Inside each item, a vertical or horizontal line is located. Participants' task is to identify the orientation of the line inside the target. Consequently, the additional singleton task is a compound search task because the features that define the target are different from the features that define the response. This allows differentiating between processes that affect visual target selection from processes associated to response selection. Target and distractors are of the same color, but in half of the trials one distractor has a different color than all the other items in the display (e.g., a red diamond among green diamonds). This distractor is called an additional color singleton distractor (CSD). The CSD is, like the target, distinct from the rest of the distractors. Whereas the target is a form/orientation singleton, the CSD is a singleton due to its unique color. Importantly, the CSD is more salient than the target and the distractors because the bottom-up activation by color is more powerful than the bottom-up activation by any other dimension (see also Wolfe, 1998a). In this example, a red diamond among green diamonds is more salient than a green circle among green diamonds. As the target is constant across the experiment (i.e., its defining feature does not change), participants should have the intention to ignore the color red. Nevertheless, response latencies were significantly delayed by 20 ms in trials in which the CSD appeared (Theeuwes, 1992). Theeuwes concluded from the results that a CSD captures attention involuntarily during preattentive processing. Salience is the single determining factor of the first deployment of attention. These results supported the view that a top-down goal (i.e., “Ignore the red item”) cannot modulate visual selection in the first stage of processing.

There is, however, one possibility for top-down control to affect search: Participants can voluntarily modify their *attentional window.* The attentional window is the area of the visual field that participants screen preattentively and for which salience values are computed. If this field is very narrow and attention very focused, attentional capture does not occur because the chances are high that the stimulus is not inside the attentional window. Yet, if the attentional window is widely distributed across the display, all of the stimuli fall inside the focus of attention and the most salient stimulus captures attention. The first empirical evidence confirming this hypothesis came from a study that manipulated the attentional window indirectly via different display types (Theeuwes, 2004). Furthermore, Belopolsky, Zwaan, Theeuwes, and Kramer (2007) manipulated the attentional window directly by asking participants either to indicate whether the items in the display formed an upward or downward pointing triangle (wide attentional window) or to identify the fixation point (focused attentional window) prior to the visual search task (for a similar procedure see Belopolsky and Theeuwes (2010)). In those studies, attentional capture did not occur in the focused attentional window condition. Theeuwes argued that parallel search is an indicator of a wide attentional window, whereas a focused attentional window requires serial search: “when search is serial (or partly serial) preattentive, processing plays no or only a minor role because due to the serial nature of the task, attention is focused on a restricted spatial area thereby circumventing preattentive processing outside that area” (Theeuwes, 2010, p. 79). Thus, the prevention of capture by a focused attentional window comes at the price of slower response latencies because preattentive processing is restricted to a very small area of the visual field. The attentional window hypothesis is also supported by a study by Lu and Han (2009), which reported a decrease or even absence of attentional capture in difficult, serial search.

The distinction between the parallel, preattentive stage of visual processing and the serial, focal stage is central to Theeuwes' account. The absence of attentional capture in a search task that involves serial search does not falsify his theory because serial search very likely is associated with a focused attentional window. Consequently, a test of his theory has to guarantee that search in the additional singleton task is completed in parallel. On the other hand, a slowing of response time by a salient distractor cannot be attributed to attentional capture if search is serial. In such a case, the delay could also be due to distraction during the serial stage of search. In this thesis, the term "attentional capture" refers specifically to distractor interference during parallel search, whereas the more general term "distractor interference" is used for a distraction effect in parallel or serial search tasks.

A variant of the additional singleton task employs variable target and distractor forms (Theeuwes, 1991). Participants have to look for the unique form among homogeneous distractors: the target is either a diamond among circles or a circle among diamonds. A CSD appears in 50% of the trials. Search in this variant of the task is more difficult than in the variant with constant target and distractor forms (cf. Theeuwes, 1992) and the average amount of attentional capture is approximately 120 to 150 ms.

Theeuwes, DeVries, and Godijn (2003) investigated whether attentional capture is reflected in eye movements. They employed two different experiments: one with a constant target (cf. Theeuwes, 1992) and one with a variable target (cf. Theeuwes, 1991). In the former experiment, eye movements were not distracted by the presence of the distractor. Participants' hit rate of the first saccade on the target did not differ between trials with a CSD and those without a CSD. In the latter, however, the eyes went as often to the target as to the CSD (the hit rate was 38% in both conditions). Theeuwes et al. (2003) termed this effect, in analogy to the results with the covert search task, *oculomotor capture*. Further evidence supporting the stimulus-driven capture account was delivered by studies that more closely examined the time course of saccadic latency. Those showed that fast saccades (i.e., saccades with a saccadic latency of less than 200 – 250 ms) are completely driven by bottom-up activation. Thus, a high proportion of them lands on the most salient item of the display. Slow saccades and any saccade that has been initiated after the first saccade, however, are under top-down control (Siebold, van Zoest, & Donk, 2011; van Zoest, Donk, & Theeuwes, 2004). These results corroborate the proposition that the early phase of overt attentional orienting cannot be controlled by task goals.

The term oculomotor capture, which is used by Theeuwes, is misleading because it implies that an automatic shift of the eyes to a distractor is not caused by attentional processing. The differentiation between attentional and oculomotor capture could therefore consolidate the obsolete view of attentional processing and oculomotor behavior as two separate systems (Findlay & Gilchrist, 2003). Theeuwes et al. (2003) based their study on the view that covert and overt attentional processing are tightly connected. Therefore, this thesis uses the term of *covert attentional capture* to refer to an automatic shift of covert attention to an irrelevant distractor, indicated by a delay in manual response latencies, and the term *overt attentional capture* to refer to an automatic shift of eye movements towards the irrelevant distractor. In both cases, attentional capture occurs only in the preattentive stage of search.

Theeuwes' studies in the early 1990s (1991, 1992) sparked a debate about the role of goal-directed attentional control in visual search. His theory provided an interesting and simple answer to the question how irrelevant information is filtered out by attentional processing: in the first phase of attentional orienting, relevance does not play a role at all and goals can only affect serial search. The first study that challenged Theeuwes' account investigated an alternative explanation for his results: Attentional capture is not the cause of stimulus-driven control in the preattentive phase, but of a top-down search mode.

2.1.4 The Search Mode Account

In 1994, Bacon and Egeth investigated the hypothesis that attentional capture in the additional singleton task is not due to the salience of the CSD, but due to a search mode of the participant that is tuned towards the detection of singletons. The specific characteristic of the additional singleton task is the presence of two singletons among homogeneous distractors. Bacon and Egeth argued that this kind of display induces participants to look for any discrepant item in the array and not for the target feature. They used two different display types in one experiment to test their hypothesis. The display in the homogeneous distractors condition employed homogeneous distractors and the two singletons (target and CSD) (cf. Theeuwes, 1992). In the heterogeneous distractors condition, however, target and CSD were not singletons because the distractors had different shapes (a diamond, a square, and a triangle). Thus, every item in the display was a singleton. As expected, response latencies were only delayed by the CSD in the condition with the homogeneous distractors, but attentional capture did not occur with heterogeneous distractors.

The authors took this result as evidence for two different search modes induced by different display types. If participants search for a singleton target among homogeneous distractors, they use *singleton search mode* and search for any discrepant item in the display. Thereby, they are very prone to capture by the CSD. On the other hand, if the display contains distractors with heterogeneous forms, singleton search mode is not a viable search strategy. In this case, participants apply *feature search mode* and look for the specific target feature. By actively using their top-down knowledge about the target feature they are not susceptible to attentional capture. Consequently, attentional capture depends on participants' search strategy and is top-down mediated.

The assumptions of the search mode account were corroborated by an influential study by Leber and Egeth (2006). The authors induced the different search modes with two different trainings. In the singleton training group, observers searched for a form target, whose shape varied from trial to trial (circle, diamond or triangle), among squares. In the feature training group, observers searched for a constant form target (a circle) among heterogeneous distractors (diamonds, triangles and squares). Both groups completed 480 training trials. After that, participants of both groups completed 480 test trials in which they searched for a circle among squares (basically the same display that was used in the study by Theeuwes (1992)) – a display which induces singleton search mode. Both in the training and in the test phase, a CSD appeared in 50% of the trials. Only the singleton training group showed significant attentional capture in the test phase. The feature display group was not distracted by the CSD, although they could have applied singleton search mode in the test phase. This study demonstrated that a search mode that one was trained in persisted with another display that would have allowed for a different, supposedly easier, search mode (see also Leber, Kawahara, & Gabari, 2009).

Theeuwes (2004) challenged the search mode account in the study in which he first proposed the attentional window hypothesis. Participants are typically induced to use feature search mode by including heterogeneous shapes in the display, so that target and CSD are not the only singletons in the display. Theeuwes argued that distractor heterogeneity fundamentally alters the search process needed to find the target. Duncan and Humphreys (1989) pointed out that visual search depends on the dissimilarity between distractors and target and on the similarity between distractors. If distractors are similar to each other, but dissimilar to the target, search is fast and parallel. Yet, if the distractors are dissimilar to the target and dissimilar to each other, search has to be conducted in serial to find the target. Based on this theory, Theeuwes argued that displays inducing feature search mode also reduce the salience of the target and of the CSD. Due to the decreased salience of the target, it can only be found with serial search. The crucial precondition for attentional capture, however, is that the CSD is more salient than the target and thereby can be detected in parallel search. The search slopes reported by Bacon and Egeth (1994) support his argumentation. These amounted to a maximum of 11.5 ms/item. As was mentioned in Chapter 2.1.1, 10 ms is considered to be the cut-off criterion to categorize search as parallel or serial. Although 11.5 ms is very close to this cut-off, Theeuwes (2004) argued that search was at least partially serial in the study by Bacon and Egeth and thereby led to the adoption of a focused attentional window. He tested this hypothesis by comparing different set sizes: one with 12 stimuli and another with 20 stimuli. The CSD was assumed to be more salient in the latter display than in the former because more homogenous distractors were included. Each of the two display types contained one triangle, one square and one diamond and either nine or 17 circles. Observers had to indicate the line orientation inside the diamond shape. In 50% of the trials, one of the circles was red (i.e., the CSD). All of the other stimuli were green. Analysis revealed that search slopes did not differ significantly from zero. Consequently, search was parallel in both conditions. The CSD prolonged response latencies in both set sizes by about 65 ms. Theeuwes concluded that this result provides unequivocal support for the attentional window hypothesis: Even though participants could have used feature search mode with the displays employed, they were distracted by the CSD.

Theeuwes (2004) noted that the distinction between feature and singleton search mode would be compatible with his account if feature search mode means to apply a focused attentional window and singleton search mode to apply a wide attentional window. His attentional window account underlines the proposition of the stimulus-driven capture account that during the preattentive, early phase of visual selection, top-down modulation is not possible. Yet, once an item has been selected, top-down strategy does play a role.

Conflicting evidence was provided by a study by Wienrich and Janczyk (2011). The authors could not replicate the central result of Theeuwes (2004) in seven independent replications. They used the same design and the same number of participants (eight to twelve), but did not observe attentional capture by the CSD. These results clearly support the search mode account because they show that attentional capture can be controlled in parallel search. So far, an account that reconciles these diverging results has not been put forward. The next and last theory on visual search that is discussed in this thesis rejects the idea of stimulus-driven capture and argues that top-down control during the preattentive phase is possible, given certain requirements. The assumptions of this theory provide an opportunity to connect the cognitive process of goal-directed visual search with motivational accounts of goal-directed behavior.

2.1.5 The Dimension Weighting Account

The dimension weighting account (Krummenacher & Müller, 2012; Krummenacher, Müller, & Heller, 2003; Müller, Heller, & Ziegler, 1995; Müller & Krummenacher, 2006a) is a modification of feature integration theory and argues that top-down knowledge guides visual search during the preattentive stage. Feature saliency maps are grouped by their respective dimensions (e.g., maps for blue, red and green are grouped in the dimension color), resulting in dimension-specific saliency maps. Before these dimensional maps are integrated into a saliency map, different attentional weights, which are determined by top-down knowledge of the target, are assigned to different dimensions and within these, to specific features. The total amount of attentional weight, which can be distributed among features and dimensions, is limited, so that an increased weight for one dimension (e.g., color), includes a decreased weight for other dimensions (e.g., orientation). For one stimulus to surpass the response threshold, the dimension in which it is defined needs higher weight than the other stimuli. Importantly, the distribution of attentional weights and the maintenance of the top-down set require effort. Thus, in contrast to the stimulus-driven capture account and feature integration theory, the dimension weighting account proposes that the saliency map already contains top-down information and therefore guides the first phase of attentional deployment to the likely target location. The main assumption of the account - that top-down information affects the computation of the saliency map - is very similar to the guided search model by Wolfe and his colleagues (Wolfe, 1994; Wolfe, Cave, & Franzel, 1989).

The dimension weighting account is central for this doctoral thesis because its authors implemented several studies with the additional singleton task in order to counter the stimulus-driven capture account (Theeuwes, 2010) and the search mode account (Bacon & Egeth, 1994). Like the search mode account, the dimension weighting account assumes that attentional capture can be modulated by top-down control, but the accounts differ on the causing factors. According to the search mode account, an observer is either susceptible to attentional capture (if she looks for any singletons) or not (if she applies feature search mode). Zehetleitner et al. (2012), however, argued that feature and singleton search mode are not two distinct attentional control modes, but are the consequences of different attentional weight settings induced by different displays and are better described as two poles of a continuum. When an observer knows that stimuli of a certain color can always be ignored (e.g., if the CSD is always red) the attentional weight for this color is decreased and its activation on the saliency map thereby reduced. Importantly, observers can only reduce a feature weight if they *practice* the suppression of the distractor feature several times. Thus, the dimension weighting account argues that participants in feature singleton mode are less prone to attentional capture not because they selectively search for the target feature, but because they practiced distractor suppression more effectively than the singleton group.

Zehetleitner et al. (2012; Exp. 2) tested this hypothesis with a modification of the study by Leber and Egeth (2006). Like in the original study, participants in feature search mode and in singleton search mode completed 480 trials in the training and test phase. Yet, in contrast to the study by Leber and Egeth, the display in the training phase did not contain a CSD. The results showed that participants both in the feature search group and in the singleton search group showed similar attentional capture effects during the test phase. This suggests that the cause for a reduction of attentional capture in the experiment of Leber and Egeth was not feature search mode per se, but distractor practice. Without experiencing the CSD during the test phase, participants in the feature search group were as distracted by it in the test phase as the participants in the singleton search group. This argumentation, however, does not explain why participants in the feature search group showed less interference than the singleton search group in the study by Leber and Egeth. According to Zehetleitner et al. (2012), participants in both groups had the opportunity to practice distractor suppression. The most probable explanation is that feature search mode depends on sufficient distractor practice: If participants are forced to search for the specific target feature and if they encounter the salient distractor while practicing this search mode, they are very effective at suppressing the salient distractor later on. This explanation, however, is challenged by another result of Zehetleitner et al. (2012; Exp. 5). Manipulating search mode within instead of between participants, the authors showed that participants in singleton search mode and in feature search mode reduced their amount of attentional capture to a non-significant level. The authors concluded that distractor practice is the key precondition for successful modulation of attentional capture.

Supporting evidence for the decisive role of distractor practice comes from a study by Müller, Geyer, Zehetleitner, and Krummenacher (2009). First, they manipulated the amount of distractor practice that participants could acquire by presenting them with a first block of trials that either contained a CSD in each trial or in no trial at all. Second, they manipulated the proportion of CSD trials: In later blocks, 20%, 50%, or 80% of trials contained a CSD. As the top-down weight setting procedure requires effort, the authors assumed that participants only apply this suppression strategy effectively if they have a high incentive to do so. If the distractor proportion is high, the incentive for participants to apply distractor suppression should be high, and vice versa. As expected, distractor interference was significantly reduced when participants encountered the distractor in 100% of the practice trials and when the probability of distractor appearance in a later trial was high (80%). In Experiment 2, participants who encountered the distractor in 100% of the practice trials and in 50% or 80% of later trials were not distracted by the salient distractor at all.

Although the assumptions of the dimension weighting account - like the assumptions of the stimulus-driven capture account (Theeuwes, 2010) and the search mode account (Bacon & Egeth, 1994) - were formulated exclusively based on studies measuring covert attention, they were also tested in one eye tracking study. Geyer, Müller, and Krummenacher (2008) investigated which attentional process drives attentional capture in the additional singleton task. They used a task similar to the one used by Müller et al. (2009) and varied the relative proportion of CSD trials (20%, 50%, or 80%). The CSD delayed response latencies only in the 20% and 50% CSD conditions, suggesting that a high incentive to suppress the distractor increased participants' top-down control in the 80% condition. In addition, the CSD prolonged saccadic latencies only in the 20% condition. In all of the conditions, a higher proportion of fixations landed on the target than on the distractor. The authors concluded that attentional capture in constant target search is mainly due to the capture of covert attention. This is in line with the results of Theeuwes et al. (2003).

Apart from the studies by Theeuwes et al. (2003) and by Geyer et al. (2008), there have not been any studies that the author knows of that investigated overt attentional capture in the additional singleton task. This issue seems to be settled for the research community and attentional capture in search for a constant target is regarded as a mainly covert effect. The differing effect of the CSD on covert and overt attention indicates that not every bias in covert attention leads to a bias in overt attention. As was pointed out in Chapter 2.1.2, the function of covert attention is to prepare eye movements. If covert attention is captured by the CSD – as indicated by a delay in response times – why does it not lead to a significant proportion of saccades on the CSD? Theeuwes et al. (2003) argued that participants in constant target search can quickly disengage covert attention from the distractor location before a saccade to this location reaches the activation threshold. On the other hand, if the CSD feature and the target feature are not known, disengagement of covert attention from the CSD requires more time and consequently leads to a saccade to the CSD. This explanation is supported by the proposition of the active vision account and related theories (Baldauf & Deubel, 2008; Findlay & Gilchrist, 2003; Findlay & Walker, 1999; Godijn & Theeuwes, 2002; Trappenberg et al., 2001). As bottom-up and top-down saccades can be programmed in parallel, a first sweep of activation to a distractor location does not automatically lead to a saccade to the same location and does not prevent the parallel programming of a target saccade. If covert attention moves fast from the distractor location to the target location, the activation for the target saccade can be increased before the distractor saccade is initiated. This suggests that the threshold for covert attention to engage is lower than for overt attention, which is not explicitly stated by these models.

Furthermore, stimulus eccentricity (i.e., the distance between the center of the screen and the center of the stimuli) was 7.25° and 9.2° in the studies by Geyer et al. (2008) and by Theeuwes et al. (2003). Although the stimuli thus were clearly positioned in the peripheral visual field, a fixation in between the stimuli and the center of the screen, at roughly 3° to 4° of eccentricity, might have been enough to perceive the target line with parafoveal vision. The stimulus eccentricity of 13.5° used by Müller et al. (2009) seems far better suited to observe possible CSD effects on eye movements.

Taken together, the studies investigating the effect of dimension weighting prove that top-down control during the preattentive stage is possible, given that at least one of two prerequisites is met: First, participants have to *practice* distractor suppression by experiencing the distractor, for example in a practice block. Second, as the process of weight-setting requires effort, observers need an *incentive to apply distractor suppression* consistently throughout the experiment. Both processes have been shown to decrease attentional capture alone and in their combination. The effect of practice was investigated in the study by Zehetleitner et al. (2012), the effect of incentive was examined in the study by Geyer et al. (2008) and the combined effect was reported in the study by Müller et al. (2009). The dimension weighting account does not specify how these two requirements relate to each other, but the results suggest that each one of them is sufficient to at least significantly reduce attentional capture. Furthermore, the results in favor of the dimension weighting account challenge the assumptions of the search mode account: attentional capture can be controlled by distractor practice even in singleton search mode (Zehetleitner et al., 2012, Exp. 5).

It is interesting to note that the dimension weighting account identifies a process as the relevant precondition for goal-directed control over attentional capture that relates to observers' motivation. It states that the attentional weight-setting procedure is determined by the effort that observers invest and that the main factor affecting effort is the incentive to apply distractor suppression. The authors do not elaborate on the meaning of the term "incentive" and do not refer to its motivational underpinnings. In the experiments, it is manipulated in an entirely cognitive way, namely as the *proportion of distractor trials* and the resulting *expectancy of the observer to encounter the distractor*. The dimension weighting account does not make a proposition whether participants consciously develop the *intention* to suppress the CSD or not. As the account assumes that a high proportion of CSD trials serve as an incentive for participants to suppress the CSD by top-down control, it presupposes that participants are intrinsically motivated to avoid the CSD in the task, but only recruit top-down control if the effort is worth it (i.e., if the CSD appears in many trials). One has to assume that the extrinsic motivation to avoid the CSD was low in the experiments conducted to test the dimension weighting account because participants were compensated with course credit, irrespective of their overall error rate or response speed.

Incentives for the application of goal-directed control also implicitly play a role in the search mode account. This account assumes that participants have a natural tendency to apply singleton search mode whenever the target is a singleton, even though it leads to distraction by the CSD. Feature search mode, however, is only applied if the display consists of non-singletons. The account does not provide an explanation for the preference of the singleton search mode. An obvious reason could be that participants weigh the effort that they have to invest in feature search mode against the benefits of it. If only one additional singleton appears in only half of the trials, the distraction caused by the CSD might be considered tolerable and the less effortful singleton search mode therefore a viable strategy. The application of singleton search mode during search among a non-singleton display, however, leads to a higher amount of distraction. In this case, participants might decide to apply the more effortful feature search mode in order to allow them to search efficiently. Therefore, one could argue that both theories assume that participants accept a minor degree of distraction by the CSD and only activate effortful goal-directed control if the amount of attentional capture reaches a certain threshold. The avoidance of a significant amount of distraction serves as the incentive to apply top-down control.

In conclusion, this means that attentional capture in the additional singleton task is actually not a purely cognitive effect, but, at least partly, also a motivational one. If incentives determine whether goal-directed control is activated, an increase in incentives should lead to a decrease in attentional capture. What if a delay of the response by 20 or 50 ms would have a direct consequence for participants' goal pursuit? What if attentional capture only arises because participants just invest that much effort that is necessary to complete the task in due time without too much errors? Such questions are a bit exotic for the aforementioned theories. Each of them conceptualizes top-down control as a purely cognitive process, but considers intentionality as the hallmark of it. The following chapter (2.2) outlines that this leads to a fragmentary concept of goal-directed search because intentions are always tied to motivational processes. An account that integrates the cognitive processes outlined by the dimension weighting account with the concept of goal-importance can deliver the basis for more thorough empirical tests of top-down control in visual search.

2.2 Goal-Directed Visual Search: the Motivational Perspective

2.2.1 The Role of Goal Importance for Goal-Directed Attentional Control

"When we speak about top-down selection we imply that selection is completely under control of the intentions of the observer. In other words, selection is completely volitional: at any time, a person can choose at will from the environment what to select." (Theeuwes, 2010, p. 77). The three aforementioned theories (see Chapters 2.1.3 - 2.1.5) agree on the fact that intentionality is the defining characteristic of goal-directed attentional control, yet none of them elaborates on the role of intentions. The theories, implicitly or explicitly, assume that there is only one intention in the lab, namely to follow the task instructions. In the case of the additional singleton task, these ask participants to find the target as fast as possible while avoiding errors.

According to the model of action phases by Heckhausen and Gollwitzer (1987), intentions are formed when the individual moves from the deliberative, predecisional phase to the implemental, postdecisional phase. In the former phase, abstract wishes and needs are transformed into concrete goals which the person intends to attain, while in the latter, the individual is committed to initiate actions in order to pursue the desired goal. A goal is a "desirable future state of affairs one intends to attain through action" (Kruglanski, 1996, p. 600). Goals are a central concept for many influential theories of self-regulation (Carver & Scheier, 1998; Gollwitzer, 1990; Kuhl, 1994; Rothermund, 2011). The "popularity" of goals can be explained with their integrative function: They serve as the link between theories of motivation which identify needs and motives as the determining factors of human behavior (Maslow, 1943; McClelland, 1985; Ryan & Deci, 2000) and cognitive theories of information-processing (Gollwitzer, Kappes, & Oettingen, 2012). Accordingly, theories of self-regulation define goals as cognitive representations within the associative memory network that mediate the effect of motivation on individuals' experience and behavior (Austin & Vancouver, 1996; Kruglanski, 1996).

Kruglanski et al. (2012) proposed a comprehensive and integrative account of motivated cognition: *cognitive energetics theory.* It is comprehensive, as it is "meant to pertain to all instances of goal-directed thinking" (p. 3), such as impression formation, perception, judgment and decision making. Furthermore, it aims to integrate effects of motivated cognition that were derived from different theories within one framework. Cognitive energetics theory argues that motivation refers to the amount of energy that is invested in goal pursuit. In more cognitive terms, energy can be referred to as the amount of mental effort that is invested. Based on Lewin's field theory (Lewin, Heider, & Heider, 1936), the main assumption of cognitive energetics theory is that motivated cognition is determined by the interplay of a *driving force* aimed at facilitating a certain activity and a *restraining force* aimed at preventing this activity. Each force has a direction (i.e., the goal which one wants to attain) and a magnitude (i.e., the desirability of the goal). The driving force is subdivided into a *potential* and an *effective driving force*. The former refers to the "maximal amount of energy the individual is capable of investing in a given goal pursuit" (p. 4) and is the product of *goal importance* and the *resource pool*, i.e. the amount of available mental resources. Like Atkinson (1964), the authors argue that goal importance is the product of the subjective *value* of the goal (i.e., its *desirability*) and the *expectancy* of goal attainment (i.e., its *feasibility*). Thus, it refers to the degree of commitment towards the goal. The effective driving force is the "actual amount of energy ultimately invested in goal pursuit" (p. 4). The restraining force, on the other hand, is determined by the additive function of the amount and importance of *alternative goals*, of the *task demands* and of the individual's inclination towards *resource conservation.* Given that the effective driving force is equal to the restraining force, the probability of goal attainment is the product of the person's skill and the magnitude of the potential driving force. If the effective driving force is lower than the restraining driving force, goal attainment is not possible. Based on the theory of motivation intensity (Brehm & Self, 1989) and the principle of resource conservation, cognitive energetics theory argues that the amount of effort invested in goal pursuit is proportionate to the amount of task demands (i.e., task difficulty). This means that individuals do not invest more effort in a task than the task requires.

Kruglanski et al. (2012) cite two studies on self-regulatory depletion as empirical evidence showing that goal-importance is a determining factor of the potential driving force. Muraven and Slessareva (2003) showed that the depletion of self-regulatory resources depends on the relevance of self-control for participants. In one group, participants had to write down each thought that came to their mind, but to suppress any thought about a white bear. They were instructed to immediately redirect their thoughts if they mistakenly thought of a white bear. This procedure leads to the depletion of self-regulatory resources (depletion condition). The other group had to memorize a list of words that were given to them by the experimenter. Afterwards, participants had to perform several problem-solving tasks (i.e., tasks that require the expenditure of self-regulatory resources). The importance of the task goal "self-control" was manipulated. Some participants received an incentive to exert self-control (they were either told that the results in the problem solving task were vital for research on Alzheimer or were given monetary incentives based on their performance), while others did not. Results showed that participants whose self-regulatory resources were depleted and who did not receive an incentive for self-control stopped working on the tasks sooner than the other groups. Among those participants who received an incentive for self-control, the depletion and the no-depletion group did not differ in their performance in the dependent variables. Thus, participants were able to compensate the depletion of their resources if they regarded the task goal as important.

A study by DeWall, Baumeister, Mead, and Vohs (2010) also delivered evidence for the role of goal importance, although it was manipulated only implicitly. The authors manipulated participants' self-regulatory depletion and their power. In the depletion condition, participants had to watch a video clip while ignoring words presented at the bottom of the screen. Participants in the no-depletion condition watched the same video clip without any specific instructions as where to direct their attention. Furthermore, participants were assigned to a high power, low power or control condition. The dependent variables were participants' performance in a dichotic listening and a problem solving task. Among the participants with high power, there was no difference between those in the depletion and those in the no-depletion group. Yet, participants who were in the low power group and whose resources were depleted showed performance costs in the tasks in comparison to low power participants whose resources were not depleted. Kruglanski et al. (2012) argued that participants in the high power condition regarded the task goal as more important and therefore were able to compensate the loss of resources.

The effect of goal importance was also studied in the realm of selective attention. In a first experiment, Vogt, De Houwer, Moors, Van Damme, and Crombez (2010) tested the hypothesis that the activation of a goal leads to the activation of associated means with which to attain these goals. Another study examined whether an attention bias towards goals depends on the importance of the goal (i.e., on its desirability and on its feasibility) (Vogt, De Houwer, & Crombez, 2011). Vogt et al. (2010) employed a modified spatial cueing task to measure selective attention. Participants had to indicate whether a black square appeared on the left or the right side of the screen. Prior to target presentation, a word was presented at both of the possible target locations. This cue could either be valid - if the target appeared at the same location as the cue - or invalid - if the target appeared at the opposite location of the cue. The trial ended with the goal task: After the target disappeared, participants were asked to press the spacebar whenever one of two goal-relevant words had been presented as a cue before the target. In a whole, eight different words were used as cues. Two of them were goal-relevant because participants were promised a financial reward if they correctly identified them in the goal task. The delay of response times in invalid trials compared to valid trials is termed spatial cueing effect. In this study, the spatial cueing effect was significantly larger for trials with goal-relevant cues compared to trials with control cues. Analysis revealed that goal-relevant cues did not speed up response latencies in valid trials, but increased response latencies in invalid trials. This result indicates that participants needed more time to disengage attention from the goal-relevant stimuli than from the control stimuli. As it was not relevant for the task to orient or engage attention to the goal-relevant cues, the authors assumed that this attentional disengagement effect is an unintentional process.

In a later experiment, Vogt et al. (2011) addressed the question of how visual selection is affected by the activation of multiple goals. They hypothesized that observers' attention is selectively allocated to those stimuli that are relevant for the goal which is more important to them. The experimental trial consisted of a dot-probe task and a goal task. In the dot-probe task, each trial started with two white rectangles being presented on the screen. After 500 ms, two words were presented in those rectangles and disappeared after 350 ms. Then, a black dot appeared in one of the two rectangles. Participants had to indicate the position of the dot. The trial continued with the goal task. In this task, a word appeared in the center of the screen and participants had to press the spacebar when they thought that this word was goal-relevant. Four words were used in the goal-task: two words were defined as goal-relevant and two words were defined as control words. Goal importance was manipulated by associating different rewards with the goal-relevant words. Participants were told that they would receive 90 points for the high-value word (e.g., boot) and 10 points for the low-value word (e.g., field). The authors tested the hypothesis that attention is biased towards the high value words. As expected, response latencies were 10 ms faster when the dot appeared at the location of a high-value word compared to when it appeared at the location of a low-value word. In a follow-up experiment, the authors applied the same procedure to words with a high (i.e., participants received a reward in 90% of the trials) and low expectancy of success (i.e., participants received a reward in 10% of the trials) and found the same pattern of results. Consequently, the results confirmed the hypothesis. Observers automatically direct their attention towards stimuli that are associated with a goal of high desirability or high feasibility.

Vogt et al. (2011) concluded that the attention bias towards goals is a stimulus-driven effect that is not mediated by conscious intentions. The orienting of attention towards the goal-relevant words was not associated with any benefit for the participants. Therefore, these studies do not provide a test of goal-directed attention. Still, the study by Vogt et al. (2011) is inspiring because it sheds light on the fact that in any given situation, and thus also in the lab, more than one goal affects cognition. The goals that are activated during an experiment should be defined by the task instructions, but a participant could also think of completely irrelevant (when relevance is defined in relation to the task goal of following the task instructions) goals, such as going shopping or meeting a friend - especially when she is not motivated to perform the task.

Experimenters implicitly assume that participants either enter the lab with intrinsic motivation to adhere to the task instructions or that otherwise extrinsic motivation (e.g., course credit or a monetary payment) can compensate for the lack of intrinsic motivation. This should inhibit the distracting influence of irrelevant goals. Even assuming that participants are motivated to follow the task instructions, these often contain several goals and one cannot be sure that participants pursue each of them with the same level of commitment. In visual search tasks, participants are normally told to find the target as fast as possible and to avoid errors. Concerning the additional singleton task, it is important to differentiate between the goal to complete the experiment, the goal to avoid errors, the goal to give fast responses and the goal to avoid distraction by the color singleton distractor (CSD). Assuming that the individual takes part because she is a bit interested in what the task is about and/or because she still needs course credit, the first three goals should be more important to her than the last. The value (i.e., the desirability) of the first goal derives from the incentive for taking part in general (e.g., course credit) and from the observer's interest in the task. The value of avoiding errors and giving fast responses derives from two sources as well: Firstly, fast responses and few errors are instrumental to the goal of leaving the lab sooner. Secondly, they satisfy the need for competence if feedback on performance is given (Ryan & Deci, 2000). The goal of avoiding the CSD, however, which is not even explicitly communicated to participants in all of the studies that were discussed in Chapter 2.1, seems to have no value to participants. The distraction by the CSD entails such a short amount of time that it is hardly perceived by the participant[1] and a delay of 20 ms does not interfere with any of the other goals. Therefore, one has to assume that the effortful control of the CSD is less important to participants, especially because they have other

---

[1] The distraction might be perceived by participants in trials in which they confuse the distractor with the target. This might occur in some trials, but the attentional capture effect has never been observed in error rates.

concrete goals on which they can focus and which are instrumental for higher-order intrinsic or extrinsic goals. Based on the principle of resource conservation (Brehm & Self, 1989; Kruglanski et al., 2012), which was described in the beginning of this chapter, participants should be motivated to invest their effort into pursuing those other goals. Even if one, in accordance with the dimension weighting account (Geyer et al., 2008; Müller et al., 2009), assumes that participants are intrinsically motivated to suppress the CSD, the question remains how strong their commitment towards this goal is.

Based on this argumentation, this doctoral thesis aims to study the hypothesis that attentional capture in the additional singleton task is, to a certain extent, a motivational effect. It is assumed that participants are reluctant to invest mental effort into goal-directed control because their natural commitment towards the suppression of the CSD is low. To test the hypothesis that top-down control modulates the preattentive phase of visual selection, one has to examine attentional capture in a setting in which participants are highly motivated to suppress the CSD. Incentives for successful goal attainment are the most convenient and effective method to manipulate the desirability, and thereby the importance of a goal, in the lab. The next chapter (2.2.2) presents the results of recent experiments showing that monetary performance-contingent incentives facilitate cognitive control and visual selection in various tasks.

2.2.2 The Enhancing Effect of Performance-Contingent Incentives on Cognitive Control and Goal-Directed Search

Recent years have seen a boost of research on the effect of motivation on visual selection and cognitive control (for a review see: Braver et al., 2014). Cognitive control is defined as "a collection of mechanisms, including perceptual selection, response biasing, and online maintenance of contextual or goal information, by which the human cognitive system adaptively configures itself to optimally perform specific tasks" (Chiew & Braver, 2011, p. 1). According to this quite broad definition, top-down control in visual search is a component of cognitive control. Despite this conceptual overlap, studies on goal-directed control in visual search seldom refer to studies on cognitive control and vice versa. Research on cognitive control developed a distinct vocabulary and methodology due to its close ties to cognitive neuroscience. Whereas research on goal-directed attention works with visual search tasks and related paradigms (e.g., the dot-probe task or the spatial cueing task), theories on cognitive control are typically tested with response conflict tasks, such as the stroop task, the flanker task or the go/no-go task (Botvinick, Cohen, & Carter, 2004). These tasks share the characteristic that participants have to override a conflicting response to execute the response that is relevant for the task goal. In the stroop task, participants have to overcome the dominant response of reading the word in order to name the ink color. In the flanker task, the central target stimulus has to be categorized while the flanking distractors have to be ignored and in the go/no-go task, participants have to respond to one set of stimuli while withholding response to other stimuli. In these tasks, cognitive control is required to overcome the response conflict. Results derived with functional magnetic resonance imaging (fMRI) indicate that cognitive control is mediated by activation in the dorsolateral prefrontal cortex (Miller & Cohen, 2001).

Müller et al. (2009) pointed out that the additional singleton task can also be conceptualized as a response conflict task because the salient signal of the CSD has to be overridden to initiate the target response. The authors argued that evidence supporting this account was put forward by a fMRI-study (Lavie & de Fockert, 2006). This study showed that attentional capture in the additional singleton task was correlated with activation in the frontal cortex and was increased when working memory resources were depleted by a high-load secondary task.

In his account of cognitive control, Braver (2012) defined what he claims to be the most effective mode of cognitive control, namely proactive cognitive control. Proactive cognitive control refers to two processes: Firstly, goal representations are triggered before they need to be implemented, due to advance knowledge that they are necessary for task fulfillment. These representations are actively maintained in working memory as long as they are useful for the task. Second, goal representations are not only continuously maintained, but also updated. The cognitive system checks whether the current goals still fit with the current task demands and whether a switch in goal representations is needed. Thus, proactive control is advantageous because it inhibits interfering information at an early stage and flexibly adapts behavior to task demands. On the other hand, proactive control is an effortful process that consumes cognitive resources. The less effortful mode of reactive control only initiates cognitive control once distracting information is detected, which means that irrelevant stimuli can interfere with performance to a greater extent. The definition of proactive cognitive control resembles the definition of top-down control by the dimension weighting account (Müller et al., 2009; Zehetleitner et al., 2012). Individuals can use it, but they need to invest effort in order to constantly maintain the weight setting procedure.

The processes of pro- and reactive cognitive control can be measured with the AX performance task (e.g., Locke & Braver, 2008). In this task, participants have to give a target response when cue “A” is followed by probe “X”, but have to give a non-target response when cue “A” is followed by probe “Y” or when cue “B” is followed by probe “X” or probe “Y”. “AX” trials occur with a probability of 70% and the remaining three trial types each make up 10% of the trials. A proactive control strategy in this task means that participants, as soon as they perceive the cue, actively maintain the information provided by the cue to prepare the target response that most likely will be the correct one. This results in performance increases in “B”-cue trials (the cue “B” is always followed by a non-target response, therefore “B” always induces the preparation of the correct response) and performance decreases in “AY” trials (the expectation induced by the cue leads to the preparation of a target response, which is incorrect). On the other hand, reactive control means that a participant only initiates the response after the probe has appeared and does not use cue information for the preparation of the response. This should result in performance benefits in “AY” trials because the probe “Y” is always associated with a non-target response and in performance detriments in “BX” trials, in which the dominant response for “X” (target) has to be overcome.

In their review on current research on the interaction between cognitive control and motivation, Braver et al. (2014) refrain from suggesting a coherent theoretical background because they aim to integrate studies from social psychology, reinforcement learning, cognitive ageing and neuroscience into a common theoretical framework. This is a very ambitious project due to the different taxonomies and definitions of motivation and cognition that these approaches use. Thus, the authors note that they regard their review mainly as a starting point for closer empirical and theoretical cooperation between the disciplines. Yet, Braver also outlines that his model of cognitive control contains a possible explanation for reward effects on mental effort and performance. The implementation of proactive control requires more effort than reactive control. As people choose the less effortful task if their choice has no further consequences (Kool, McGuire, Rosen, & Botvinick, 2010), a proactive control strategy is only applied if its benefits outweigh its costs. This is in line with motivation intensity theory (Brehm & Self, 1989) which posits that people only invest the amount of effort hat is necessary to meet the task demands. Thus, an obvious explanation for the effect of reward on cognitive control is that reward for successful goal attainment increases the desirability and the importance of the goal and thereby increases the mental effort that is invested to attain the goal.

Several studies have shown that performance-contingent rewards increased performance in the AX-task by facilitating specifically proactive control strategies (Chiew & Braver, 2013; Fröber & Dreisbach, 2014; Locke & Braver, 2008). In a similar vein, performance-contingent reward was found to reduce task-switching costs (Savine, Beck, Edwards, Chiew, & Braver, 2010; Savine & Braver, 2010). In these studies, a cue informed participants which of two tasks they had to perform in the upcoming trial (“Attend Face” or “Attend Word”). After the cue presentation, a word was superimposed on a male or female face. In “Attend Face” trials participants had to decide whether the face was male or female and in “Attend Word” trials participants had to indicate whether the word on the face consisted of two syllables or more. Thus, the cue allowed participants to apply a proactive control strategy and to prepare the focusing of attention towards the task-relevant stimulus. Results showed that participants were faster to switch between the two tasks without making more errors when they were rewarded. In a cued stroop task, stroop interference was reduced in reward trials as indicated by fewer errors (Veling & Aarts, 2010). This study also suggests an effect of reward on proactive control because optimal performance depends on the extent to which participants use the task cue to prepare the correct response.

Three studies investigated the effect of performance-contingent reward on visual selection, and two of them used visual search tasks. These studies deliver different causal explanations for reward effects. Engelmann and Pessoa (2007) combined a modified spatial cueing task with performance-contingent rewards. Participants could win or lose a low or high amount of money if they detected the target fast and correctly. The target was a faint red dot superimposed either on a picture of a face or of a house. Detection sensitivity (d') for the target increased linearly with the magnitude of reward in valid and invalid trials. The authors attributed this effect to a facilitation of stimulus-driven attention by reward. They did not deliver an explanation for this facilitation effect, but they obviously assumed that reward information associated with a target strengthens the bottom-up signal of the target features. Yet, this could also be regarded as a top-down effect because the modulation of a bottom-up signal requires top-down effortful control according to the dimension weighting account (Müller et al., 2009; Zehetleitner et al., 2012).

Kiss, Driver, and Eimer (2009) studied reward effects on visual search using a pop-out search task. Reward was not associated with performance per se, but with specific target colors. The display consisted of a colored singleton target among homogenous, colored distractors. All of the stimuli were diamonds with a notch at the top or at the bottom. Participants were told to search for the target and to report the location of its notch. Importantly, the target color changed randomly between red and green and participants were told that they would receive 5 bonus points for a fast and correct detection of a high-reward color target, and 1 bonus point for a fast and correct response for the other target. As this task was a pop-out search task, the target was easily found with parallel search because it was the most salient item in the display. Therefore, successful search did not require top-down control or effort. Instead, the less effortful strategy for participants was to direct attention to the item with the highest bottom-up activation. Analysis revealed that a response for a high reward target was significantly faster than that for a low reward target.

In addition, the study also recorded EEG waves. The N2pc component, an EEG indicator of visual target selection, occurred earlier and with larger amplitude for high reward targets compared to low reward targets. As the N2pc component emerges 180 ms to 220 ms after target onset, the authors concluded that reward exerts its influence on visual target selection very early in time. The authors conceptualized this reward effect as a third component of attentional control next to top-down and bottom-up control. Yet, the fact that top-down control was not needed in this task does not preclude the possibility that intentional processes of the participants mediated the reward effects.

A very interesting study on goal-directed visual search has been conducted by Voss (2004). It differs from the two previously described studies because it provides a more thorough explanation of reward effects. Participants completed a quite difficult visual search task that probably induced serial search. The valence of the target was manipulated: one type of target signaled the potential loss of 10 points in each trial (participants received monetary reward based on their total amount of points won in the search task), while another, neutral target was not associated with loss. One group of participants had control over the danger to lose points by successfully completing an additional task that directly followed the visual search task. The other group of participants did not have control over the potential loss because they were told that it was determined by chance whether they lost the points or not (in fact, the losses were matched to a participant in the control group, so that both groups did not differ in their total amount of losses). Participants in the control group were faster in detecting the danger target and made fewer errors than the no control group. Voss explains this result with reference to action control and self-regulation: Participants who can control the negative consequence of the danger stimulus find it fast and correctly because they know that their performance will determine the outcome of the trial. On the other hand, participants whose actions cannot determine the outcome of a danger trial try to avoid the stimulus. Although this theorizing was specifically applied to attention towards negative stimuli, it can be transferred to the role of goal importance for top-down control. In the control group, the goal of finding the danger stimulus fast and correctly was very important because observers could prevent a monetary loss with this action. In the no control group however, the goal of finding the danger stimulus was not important because they could not control its consequences, irrespective of how fast they found it.

To conclude, several studies reported that performance-contingent reward enhances cognitive control and visual selection, yet their theoretical explanations differ considerably. Whereas one group of studies argues for a top-down mediation of reward effects (Chiew & Braver, 2013; Fröber & Dreisbach, 2014; Locke & Braver, 2008; Savine et al., 2010; Savine & Braver, 2010; Veling & Aarts, 2010; Voss, 2004), others suggest that reward affects cognition via bottom-up control (Engelmann & Pessoa, 2007; Kiss et al., 2009). Table 1 summarizes key characteristics of the payoff scheme of the studies cited in this chapter. This overview illustrates that the studies also differ markedly in their payoff scheme. In the studies by Braver and colleagues, participants were deceived: They were informed that they could earn a bonus monetary reward - in addition to a flat monetary fee for participation - for good performance in the experiment. Contrary to what they were told, all of the participants received the same amount of reward, regardless of their performance. The results of these studies also demonstrate that relatively low monetary rewards - the flat monetary fee was always higher than the performance-contingent bonus - have a significant effect on performance. Rewards were either monetary or abstract points that were converted to a monetary amount. The amount of reward associated with a correct and fast performance ranged from 5 ct to 50 ct.

In most studies, participants were told that rewards were dependent on performance in each trial, but many studies did not contain feedback on the exact amount of reward won in a trial. Participants in the studies by Engelmann and Pessoa (2007) and by Kiss et al. (2009) did not receive any feedback at all after the trials – this procedure may have been motivated by the experimenters' assumption that reward effects are not mediated by top-down control and therefore need not be consciously perceived. Furthermore, feedback on the accumulated amount of reward was given only at the end of blocks or at the end of the experiment in most studies. The study by Voss (2004) is the only one in which participants were informed after each trial how much points they had won so far.

In all but one study (Engelmann & Pessoa, 2007) the response latency cut-off used to define a fast response was computed based on the individual distribution of response latencies. Most studies computed the cut-off based from response latencies in the baseline block. This controls for inter-individual differences in response latencies, but not for individual learning effects across the experiment. The method employed by Voss (2004) is superior: the cut-off was computed from the last six responses and therefore adaptive to practice effects.

If the aim of the payoff scheme is to increase the importance of the conscious goal to suppress the distractor/to find the target, the association between goal attainment and reward should be clearly communicated to participants. This means that they should receive exact feedback on the amount of reward won in a trial and on the accumulated amount of reward won so far.

Table 1. *Overview of Payoff Schemes in Selected Studies on Cognitive Control and Visual Search.*

Study	Deception of participants regarding the payoff scheme	Payoff scheme communicated to participants	Response latency cut-off	Feedback on amount of reward won in a trial	Feedback on accumulated amount of reward won in the course of the experiment	Total amount of reward
Chiew and Braver (2013)	Yes	Unspecified amount of bonus for each correct and fast response; +5 $ maximal additional bonus overall	Individualized: 30th percentile (from fast to slowest) of correct baseline response latencies	Unspecific: Bonus yes/no	Feedback on total amount of reward at the end of the experiment	15$
Engelmann and Pessoa (2007)	No	50% chance of winning 1$/4$ in reward blocks and of avoiding a loss of 0.5$/2$ in penalty blocks in case of good accuracy (at least 64%) and fast responses	General: 605 ms (mean response latency + 2 *SD*s in pilot study)	No Feedback	Feedback on the accumulated amount of reward at the end of each block	Individual amount
Fröber and Dreisbach (2014)	No	+5ct for each correct and fast response	Individualized: 30th percentile (from fast to slowest) of correct baseline response latencies	Feedback on exact amount of reward	Feedback on the accumulated amount of reward at the end of the first block and at the end of the experiment	Individual amount
Locke and Braver (2008)	Yes	+25ct for each correct and fast response; -3$ for each error	Individualized: Median of response latencies in baseline blocks	Feedback on exact amount of reward	Feedback on the total amount of reward at the end of the experiment	25$

Table 1. *(continued).*

Study	Deception of participants regarding the payoff scheme	Payoff scheme communicated to participants	Response latency cut-off	Feedback on amount of reward won in a trial	Feedback on accumulated amount of reward won in the course of the experiment	Total amount of reward
Savine and Braver (2010)	Yes	Unspecified amount of bonus for each correct and fast response; +10$ additional bonus overall	Individualized: 30th percentile (from fast to slowest) of correct baseline response latency	Unspecific: Bonus yes / no	Feedback on the total amount of reward at the end of the experiment	45$
Veling and Aarts (2010)	No	+1ct/+50ct for each correct and fast response	Not reported	No Feedback	Feedback on the total amount of reward at the end of the experiment	Individual amount
Voss (2004)	No	+5ct for each correct and fast response; -10ct for each error in danger trials; 50ct for each 50 points overall	Individualized: 66th percentile (from fast to slowest) of the last six responses	Feedback on exact amount of reward	Feedback on the accumulated amount of reward after each trial	Individual amount

The studies by Kiss et al. (2009) and Voss (2004) are the only ones which examined performance-contingent reward effects on visual search, and the latter covered this topic only indirectly (the focus was the effect of control on the perception of danger stimuli). The pop-out search task that Kiss et al. used did not allow making an inference whether the reward effect was due to stimulus-driven or goal-directed selection. As goal-directed control was not needed to find the target in parallel search, it is very probable that only stimulus-driven processes were at play. Unlike the additional singleton task, the pop-out search task does not entail a response conflict because a dominant response does not need to be overridden. The study by Voss is very different from the additional singleton task, too, because it induced serial search. However, its results support the role of goal importance for top-down control in visual search. Consequently, the question remains whether performance-contingent rewards also enhance goal-directed attentional control during parallel search in the additional singleton task and thereby decrease the amount of attentional capture. Another question is how eye movements are affected by rewards – this issue has not been addressed yet.

Finally, most of the studies discussed in this chapter did not measure mental effort. For some (Engelmann & Pessoa, 2007; Kiss et al., 2009) this was not an interesting variable because they did not include mental effort in their theoretical background. Yet, according to the account by Braver (2012), mental effort is the central variable that should be affected by reward. The next chapter (2.2.3) presents studies that demonstrated that pupil dilation is a reliable and sensitive measure of mental effort.

2.2.3 Pupil Dilation as an Indicator of Mental Effort and Motivation

In the 1960s, the first studies showed that pupil dilation is positively correlated with memory load (Beatty & Kahneman, 1966) and task difficulty (Hess & Polt, 1964). Since then, other studies have gathered additional evidence that pupil dilation is a sensitive indicator of the mental effort (sometimes also termed cognitive effort) that participants invest in a task: the higher the mental effort, the higher the pupil dilation (for a review see: Beatty, 1982). More recent studies have shown that the *phasic* increase in pupil diameter and cognitive effort are positively correlated in various tasks such as language processing (Hyönä, Tommola, & Alaja, 1995), lexical decision (Kuchinke, Võ, Hofmann, & Jacobs, 2007), face processing (Goldinger, He, & Papesh, 2009), or visual search (Porter, Troscianko, & Gilchrist, 2007). Pupil dilation that is caused by mental effort can reach a maximum of 0.5 mm and the average pupil size in standard lighting conditions is about 3 mm (Laeng, Sirois, & Gredebäck, 2012).

Phasic pupil dilation is defined as the difference between the maximum pupil diameter during the relevant task and the baseline pupil diameter recorded prior to the task and is the most widely used pupillometric measure in recent research. It is sometimes also termed stimulus-related pupil dilation and indicates *transient* mental effort. *Tonic pupil dilation*, on the other hand, refers to global changes in pupil diameter defined as the difference in baseline pupil diameter between different conditions over a longer period of time and is an indicator of *sustained* mental effort (Heitz, Schrock, Payne, & Engle, 2008). Pupillometry is a useful method in research on motivated attention because it is an unobtrusive and implicit measure to test whether participants invest more effort in a cognitive task if their performance is rewarded.

Bijleveld, Custers, and Aarts (2009) tested the hypothesis that reward leads to an increase in mental effort only under high task demands. In a digit retention task, participants had to recall either three (low task demand) or five digits (high task demand) and were cued with either a low (1 ct) or a high (50 ct) reward before the presentation of the digits. As expected, reward increased phasic pupil dilation, but only in the high demand condition.

On the other hand, Heitz et al. (2008) showed that monetary incentives did not have an effect on phasic pupil dilation during a memory task, but increased the baseline pupil diameter (tonic pupil dilation) in the incentive condition compared to the other conditions. In a similar vein, in a study by Wykowska, Anderl, Schubö, and Hommel (2013) tonic pupil dilation correlated with an action-congruency effect measured in response latencies. Participants were faster to find a target in a visual search task if they had prepared a manual action for which the target dimension was relevant. At the beginning of each trial, participants were cued with two different pictures that signaled to them whether they should prepare a pointing or a grasping movement. In the search task, participants had to indicate whether a target (defined by its larger size) was present or absent. After the search task, a cue indicated which of three paper cups, positioned below the computer screen, was the target of the action. Participants either had to point at or to grasp the target cup. As size is relevant for grasping, but not for pointing, trials with a grasping cue were congruent trials, whereas trials with a pointing cue were incongruent trials. Response latencies in visual search were faster in congruent than in incongruent trials, which is referred to as the action-perception congruency effect. The change in tonic pupil size from the first to the second experimental block was taken as an indicator of the maintenance of motivation during the experiment. Importantly, the more tonic pupil size decreased from block 1 to block 2, the smaller was the congruency effect. The authors took this as evidence that the link between action plans and attention is modulated by participants' consistent motivation to prepare the movement during the experiment.

Chiew and Braver (2013) measured task performance (with response latencies and error rates) and mental effort (with phasic and tonic pupil dilation) during the AX-task in which participants could earn performance-contingent reward. They manipulated reward both block-wise and trial-wise by comparing a baseline block, in which participants could not win rewards, with a reward block, in which participants could earn rewards in some trials. Reward increased task performance and the application of proactive control both in the reward block compared to the baseline block and, within the reward block, in reward trials compared to neutral trials. This was paralleled by an increase of phasic and tonic pupil dilation in reward conditions.

Whether tonic or phasic pupil dilation is the more adequate indicator of mental effort and motivation depends on the experimental procedure and whether one wants to investigate short-term or long-term changes (Laeng et al., 2012). If reward is manipulated block-wise, both tonic and phasic pupil dilation can be adequate variables to measure mental effort. If many reward trials follow after another, tonic pupil dilation might increase as a sign of a higher baseline level of effort (cf. Heitz et al., 2008). If, however, reward is manipulated trial-wise, only phasic pupil dilation can track the transient changes in effort that are assumed to be induced by the expectancy of reward (cf. Bijleveld et al., 2009).

2.2.4 Conclusion

The previous chapters have outlined a theoretical background for a “new” concept of goal-directed attentional control which acknowledges that it is also determined by the importance of the goal that is pursued. The accounts on attentional capture discussed in Chapter 2.1 did not address participants’ motivation to find the target and to suppress the CSD and therefore only studied a fragmentary concept of goal-directed search. If the activation of effortful goal-directed control depends on the incentives for goal achievement, monetary performance-contingent incentives should decrease the amount of attentional capture. If observers can win money for correct and fast responses, they should be motivated to invest effort into the down-weighting of the CSD color in order to avoid distraction that would cost time. The testing of this hypothesis is the central issue of most of the studies conducted for this doctoral thesis.

The literature on the additional singleton task discussed in the previous chapters also shows that a relatively simple search task, in which one circle has to be found among diamonds, is affected by a lot of variables (e.g., practice with the CSD, search mode and attentional window of the participant, similarity between distractors, similarity between distractors and target, incentive to suppress the CSD). The general assumption that a CSD captures attention does not hold. Given the complexity of attentional capture with neutral, “meaningless” stimuli, it might seem surprising that the visual search task has been employed by a multitude of studies arguing for such a very simple statement. Negative emotions are assumed to capture attention because their detection is relevant for the satisfaction of one of the central human motives: the motive to survive (Öhman et al., 2001). Emotion is assumed to distract visual processing during the preattentive phase of search and to be uncontrollable by the goals of the observer. This hypothesis contradicts the basic model of visual search, which assumes that only simple visual features can guide attention during the preattentive stage (see Chapter 2.1). The next chapter (2.3) discusses the theoretical background of the *emotional attentional capture* account and the validity of empirical studies that were conducted to test it.

2.3 The Emotional Attentional Capture Account

Hardly any other attention task has been used so widely by psychologists studying the interface between selective attention and emotion and motivation than the visual search task with emotional faces. Hansen and Hansen (1988) were the first to employ this task in a seminal and highly cited study. They used the classic version of the visual search task in which the observer has to search the display and report whether the target is absent or present (Treisman & Gelade, 1980). Instead of simple objects, like circles and squares, the authors used photographs of emotional faces as stimuli. Firstly, they assumed that a threatening face is detected faster among a crowd of happy faces than vice versa. Secondly, they hypothesized that the detection of threat is so important to the individual that a threatening face is found during the preattentive stage of processing. Thus, search for a threatening face among happy faces should be parallel, whereas finding a happy face among threatening faces should require serial search. The authors did not provide an elaborate theoretical framework for the *threat superiority* (also called anger-superiority) hypothesis and only broadly referred to studies showing that emotional expressions are perceived very efficiently. Yet, the results confirmed their hypotheses and sparked many more studies on this issue (cf. Becker, Anderson, Mortensen, Neufeld, & Neel, 2011), although the study suffered from a methodological confound: The picture of the threatening face was more salient than the picture of the happy face because the former contained a black spot in one corner and the latter did not. Another study used the same design and the same stimuli as the study by Hansen and Hansen (1988), but without the dark spots on the threatening face, and did not observe the threat superiority effect (Purcell, Stewart, & Skov, 1996).

Öhman et al. (2001) proposed a more elaborate theoretical framework for the threat superiority effect and addressed the issue of visual confounds with emotional stimuli. Firstly, they argued that efficient threat detection is the result of cognitive processes executed by an evolutionary evolved *fear module* (Öhman & Mineka, 2001). This has developed out of “the need of effective predatory defense systems” (p. 486). Its function is the fast and reliable detection of predators in order to elicit a life-saving fight-or-flight reaction in time. Like Hansen and Hansen

(1988), the authors assumed that threat automatically distracts visual processing during the preattentive stage. Thus, emotional valence is considered to be a perceptual dimension like color, motion or size. This assumption strongly contradicts cognitive theories of visual search, which argue that only basic visual features can be perceived preattentively (Horstmann, Becker, Bergmann, & Burghaus, 2010; Wolfe & Horowitz, 2004). Secondly, Öhman et al. (2001) addressed the visual confound in the stimuli by Hansen and Hansen (1988) by using schematic stimuli. These are impoverished in comparison to photographs, but have the advantage that confounds of visual features can be easily controlled. The results revealed that both threatening and happy target faces among neutral faces were found with parallel search (search slopes were 3 ms/item for threatening targets and 4 ms/item for happy targets) and that threatening targets were found faster than happy targets (Exp. 2). The authors interpreted this finding as empirical support for the threat superiority hypothesis. Yet, the detection advantage for threatening faces was also observed when faces were presented upside down (Exp. 4). As the inversion of faces hinders facial processing (Farah, Tanaka, & Drain, 1995; Maurer, Le Grand, & Mondloch, 2002), this result strongly suggests that the reported threat superiority effect was not driven by the detection of threat, but by a visual confound (Becker et al., 2011). Furthermore, the authors did not provide an explanation for the result that happy faces were detected preattentively, as well. According to their account, the function of the fear module is to specifically detect threat and not each emotional expression. Still, their study suggested that happy and angry faces can be detected preattentively, which is an interesting result.

Two other studies, however, showed that search for happy and threatening facial expressions is serial. Using even more impoverished schematic faces than Öhman et al. (2001), it was observed that the search slopes for threatening faces were shallower than those for happy faces, but that both significantly differed from zero (Eastwood, Smilek, & Merikle, 2001; Fox et al., 2000).

More recent studies using photographic stimuli also demonstrated that search among faces is serial and did not report evidence for a threat superiority effect. Calvo, Nummenmaa, and Avero (2008) reported that happy, disgusting and surprised expressions are more easily detected among crowds than threatening, fearful and sad expressions. M. A. Williams, Moss, Bradshaw, and Mattingley (2005) observed a happiness and threat superiority effect. In both studies, search was serial for all of the expressions. Another study revealed evidence for a happiness superiority effect, yet set size was not varied so that a conclusion regarding parallel or serial search could not be drawn (Juth, Lundqvist, Karlsson, & Öhman, 2005).

To conclude, the results on the threat superiority effect are highly inconsistent, firstly, regarding the question of whether threatening expressions are detected faster than others and secondly, regarding the question of whether this detection advantage arises during parallel or only during serial search.

Becker et al. (2011) provided a thorough review of previous studies on the threat superiority effect. The authors concluded that the majority of the previous studies did not sufficiently control for visual confounds of their stimuli. They cite one of the reviewers of their study to illustrate the negative point of view that they have on research on this topic: "The literature on visual search using face stimuli is a morass where the bold should fear to tread. Instead, the allure of faces, emotion, and evolutionary psychology continues to attract researchers like moths to the proverbial flame." (p. 657). If photographic stimuli are used, the happy expression (and, to a lesser extent, also the angry expression) is often accompanied by the exposure of white teeth (e.g., Juth et al., 2005), whereas a neutral face normally has a closed mouth. Thus, the feature contrast of the happy face is much higher and it is therefore more salient than the neutral face (Calvo & Nummenmaa, 2008). A more subtle visual confound can arise if the same target and distractor identities are used throughout the experiment. This is always the case in studies working with schematic faces because these do not portray inter-individual differences. This allows participants to identify a single feature that distinguishes the target from its distractors (e.g., schematic faces with emotional expressions have curved eyebrows, whereas the neutral face has flat eyebrows) and to search for this feature and not for the emotion per se. Thus, schematic faces also suffer from visual confounds. Furthermore, schematic faces possess very little external validity.

A practical method to control for visual confounds in photographic stimuli is to include a condition with inverted faces. Face inversion hinders facial processing (Farah et al., 1995; Maurer et al., 2002) and thereby the perception of emotional expressions. If an emotional expression is detected faster both in upright and in inverted position, the detection advantage is, at least partly, also due to its visual saliency. If, however, a detection advantage occurs only with upright faces, it must be due to the emotional content of the face.

Becker et al. (2011) used computer-generated facial prototypes that carried happy or threatening expressions without an open mouth and exposed teeth. Participants were instructed to look for a particular target expression (e.g., "Look for an angry face"). The results showed that happy faces were found faster than angry faces. For both target types, the search slopes differed significantly from zero, suggesting that search was at least partly serial (the authors did not provide information on the size of the search slopes). Thus, the authors concluded that happy faces are detected faster than angry faces, but that none of the emotional facial expressions captures attention preattentively. The study by Becker et al. delivers more valid results than the studies that were discussed in this chapter so far because it used photographic faces and controlled for possible low level confounds. Yet, the authors did not address the association between the task goal of the observers and the emotional content of the faces.

Another important critique that has been leveled at the bulk of previous research on attentional capture by emotional faces argues that the role of the top-down task set has been neglected by many studies. Hodsoll, Viding, and Lavie (2011) argued: "However, in all previous suggestions that emotional faces may capture attention during visual search, emotion was always in some way relevant to the task, either because emotional faces appeared in the same location as the task stimuli or because they were an inherent part of the task at hand (e.g., participants were instructed to search for an emotional face or an 'odd face out' that was defined as being such by its emotional expression). […] Thus, although it is clear that emotional expression could guide the search more efficiently than a neutral expression, it is not clear that emotion is capable of capturing attention when it is irrelevant to the task." (p. 347). This critique applies to all of the studies on the threat superiority effect that were cited in this chapter so far. These either asked participants to search for the discrepant face or for one specific emotional expression. In both cases, emotion is relevant for the task set because it defines the target: the target either carries an emotional expression or not (if it is defined as the discrepant item), or it carries a specific emotional expression. Therefore, the capture of attention by emotions could be the result of a top-down goal to search for a specific emotion.

Hodsoll et al. (2011) refrained from assuming that one emotion (e.g., threat or happiness) captures attention more than other emotions and instead examined a general attentional bias towards emotional stimuli. To determine whether any emotion captures attention involuntarily, one has to define task instructions in such a way that the definition of an emotional expression is not part of the top-down task set. Therefore, the authors chose the additional singleton task as a more adequate visual search task to test the emotional attentional capture effect. In their study, participants had to find the male face among two female faces and had to indicate the orientation of the target face (i.e, whether the face was tilted to the left or to the right). The stimuli were photographs of emotional expressions. In two thirds of the trials, all of the three faces were neutral in expression. In one third of the trials, one of the three faces was an emotional face, and in half of these trials (1/6 of the total number of trials), the emotional face was one of the female distractors (i.e., the emotional singleton distractor) and in the other half (1/6 of the total number of trials) it was the male target. This means that the defining feature of the singleton distractor (i.e., emotion) was also a target feature in some trials. The authors chose this procedure, which marks a slight departure from the additional singleton task, in order to prevent that participants could implicitly infer the target position from the emotional expression on a distractor. Analyses revealed that fearful, angry and happy faces delayed response latencies significantly (Exps. 1 - 3). Furthermore, a happy target was detected faster than a neutral target (Exp. 3). The emotional capture affect was also found with female targets and angry, male distractors (Exp. 4).

These results certainly provide more thorough evidence for involuntary emotional capture than the results gathered with search tasks that either asked participants to search for one specific emotion or for the discrepant face (e.g., Becker et al., 2011; Calvo et al., 2008; Eastwood et al., 2001; Fox et al., 2000; Hansen & Hansen, 1988; Öhman et al., 2001). In the fifth experiment, Hodsoll et al. (2011) aimed to tackle an alternative explanation for their results: the emotional singleton may have popped out from the display because it was the “odd-one-out” and was different in its expression from the other two stimuli. They addressed this issue by using neutral faces as irrelevant distractors among emotional faces. A neutral singleton among angry faces did not delay responses. This finding proves that attention was not captured by the odd-one-out in Experiments 1 - 4. It does not, however, rule out visual saliency effects that could arise from the exposed teeth in the happy face or the curved eyebrows in the angry face. A salient feature on the emotional face, not the emotional expression, could have captured attention. This explanation would be in line with the results from Experiment 5 because these also showed that response latencies were significantly delayed for a neutral target among angry distractors compared to response latencies for an angry target among angry distractors. Face inversion would have been a viable option to control for possible visual confounds, but was not implemented. Consequently, the authors did not provide unequivocal evidence for emotional attentional capture.

A similar study to Hodsoll et al. (2011) was published nearly at the same time and therefore did not refer to the former study. Huang, Chang, and Chen (2011) named the same motivation for their study, namely that emotional content of the faces was always relevant for the task goal in previously published visual search tasks. Consequently, the results of those studies did not provide unequivocal evidence that emotion involuntarily captures attention. In their own study, the authors used a target attribute that was clearly not associated to any emotional content: A black dot was either positioned on the right or on the left side of one of the faces and participants had to indicate the position of this dot. In contrast to the study by Hodsoll et al. (2011), the stimuli consisted of schematic faces. Five trial types were used: a neutral baseline condition, in which all of the faces were neutral, a happy target condition, an angry target condition, a happy distractor condition and an angry distractor condition. In the last two conditions, the target was always a neutral face and one of the remaining seven faces was an emotional one. Response latencies were slower in the angry distractor condition than in the baseline condition and were faster in the angry target condition than in the baseline condition.

Response latencies in the happy target and in the happy distractor condition did not differ significantly from the baseline condition.

In Experiment 2, a condition with inverted faces was added to the design to check whether the emotional capture and the emotional facilitation effect observed in Experiment 1 were due to visual confounds. Only the baseline and the angry target and angry distractor condition were employed. The capture effect was present both for upright and inverted faces, yet was significantly reduced by 20 ms in the latter condition. The facilitation effect of the emotional target, however, was of the same size both with inverted and upright faces. The authors neglected this aspect of their analysis in the discussion and instead concentrated on the significant reduction of the interference effect in the inverted angry distractor condition. This pattern of results, however, is clearly inconclusive with regard to the effect of the angry expression. Firstly, the fact that responses to inverted angry faces also differed significantly from the baseline shows that a part of the facilitation and of the capture effect is due to the visual saliency of the emotional expression. Secondly, the facilitation effect of the angry target seems to be exclusively driven by visual saliency because it did not significantly differ between upright and inverted targets. This, of course, leaves open the question of why an angry distractor captured attention due to its emotional content, whereas an angry target captured attention due to its saliency. The angry target differed from the angry distractor in one important attribute, namely the black dot. The most reasonable explanation for the discrepancy is that the combination of the black dot, the downward curved mouth and the eyebrows was more salient (in the inverted and in the upright version) than the combination of the black dot and the neutral face. As the study employed schematic faces, such differences in feature combinations can play a decisive role because participants can learn to focus on these if they are identical in each trial. To rule out such confounds, several facial identities should be presented throughout the experiment, so that participants cannot learn one single feature or feature combination that differentiates the target from the distractors (cf. Becker et al., 2011).

Like the majority of the studies published on the threat superiority effect so far, Hodsoll et al. (2011) and Huang et al. (2011) only measured response latencies and accuracy as indicators of covert attention. Conflicting evidence for the hypothesis of emotional capture comes from an eye tracking study with schematic faces by Hunt, Cooper, Hungr, and Kingstone (2007). This demonstrated that an emotional distractor only captured attention when it was relevant for the top-down task set. In Experiment 1, participants had to fixate either a happy target or an angry target that was either presented upright or upside down. At the beginning of the trial, all of the stimuli were neutral. After 1000 ms, one of the stimuli turned happy or angry and, in the distractor present condition, one of the distractors turned angry (in the happy target condition) or happy (in the angry target condition). Saccadic latency was delayed by the distractor, especially among upright faces. 26% of the first saccades were directed to the distractor, yet this effect was not modulated by the emotional content or the orientation of the distractor or the target. Finally, analysis revealed that the eyes did not fixate the emotional distractors for a longer time than other distractors. These results identified saccadic reaction time as the only indicator of overt attentional capture that was affected by the emotional content of distractors.

In Experiment 2, participants had to fixate an upright neutral target among inverted faces or an inverted neutral target among upright faces. Consequently, the target was defined by its orientation and emotion was completely irrelevant for target identification. Like in Experiment 1, the trial started with six identical faces (either upright or inverted). After 1000 ms, one of the faces (the target) switched its orientation and in half of the trials, another face switched to an emotional one. Distractor emotion did not have an effect on saccadic latency. A happy distractor was fixated more often and for a longer period of time than an angry distractor, yet this was also true for inverted faces and therefore must have been due to the increased visual saliency of the happy face. Therefore, the authors concluded that an emotional distractor captures attention if emotion is relevant to find the target ("Fixate the happy/angry face.", Experiment 1), but not if it is irrelevant to find the target ("Fixate the upright/inverted neutral face.", Experiment 2).

The study by Hunt et al. (2007) differs in one important aspect from the studies by Hodsoll et al. (2011) and Huang et al. (2011). Target and emotional distractors were presented only 1000 ms after display onset. The abrupt onset of the target and of the emotional distractors increases their salience even further in comparison to the neutral distractors because attentional capture by abrupt onsets is regarded to be more profound than that by salient, static singletons (Schreij, Owens, & Theeuwes, 2008): Motion leads to higher bottom-up activation than color or orientation. This fact strengthens the results from Experiment 2 because it shows that the emotional distractors contained sufficient bottom-up salience. Consequently, the use of abrupt onsets cannot explain the discrepancy in results. Like in the study by Huang et al. (2011), schematic faces were used, which certainly reduces the external validity of the results and which enables participants to search for the target by the identification of simple feature differences. Yet, in contrast to the study by Hodsoll et al. (2011), and like in the study by Huang et al. (2011), an inverted faces condition was employed to rule out visual confounds.

To conclude, the results of Hunt et al. (2007) provide strong evidence against an emotional capture account, at least with schematic faces and with eye movements as the dependent variable, whereas Hodsoll et al. (2011) delivered evidence in favor of an emotional capture account, with photographic stimuli and manual response as the dependent variable. A closer look at the task goals employed in these studies reveals an explanation for the discrepant results: Whereas the target was defined as a black dot or as a combination of simple features in two studies (Huang et al., 2011; Hunt et al., 2007), it was defined by face gender in the other study (Hodsoll et al., 2011). It is important to keep in mind that attentional capture refers to the distraction of the first orientation of attention during the preattentive phase (see Chapter 2.1). Thus, Hodsoll et al. (2011) assumed that searching for a face of a certain gender among distractor faces of the opposite gender can be executed in parallel. This is a contentious assumption: Wolfe and Horowitz (2004) argued in their review that the empirical evidence for preattentive guidance of emotions and of faces in general is weak. The same criticism that was issued at the hypothesis of emotional valence as a perceptual category applies to gender (cf. Horstmann et al., 2010). Hodsoll et al. (2011) did not vary set size to test whether search was parallel (this was not done by Hunt et al. (2007) and Huang et al. (2011), either), so it could be that the distraction effect by the emotional face occurred during a serial search that required focusing shortly on each face before moving on. If that was true, it would provide another explanation of the assumed capture effect of emotion. Once the emotional distractor was fixated during serial search, the eyes fixated it longer than a neutral distractor. Thus, the emotion might have delayed the disengagement of attention (Belopolsky et al., 2011; Fox et al., 2002). This would certainly be an emotional distraction effect on attention, but it would not be emotional attentional capture. The delayed disengagement account could not be ruled out by Hodsoll et al. (2011) because they did not monitor eye movements. Thus, it could be in line with their results that participants had to look at the pictures one by one to report where the male or female target is.

Interestingly, Hodsoll et al. (2011) did not address the results of Hunt et al. (2007), although these were published four years earlier. Probably they dismissed this study as irrelevant because of the use of abrupt onsets and the monitoring of eye movements. In conclusion, the current state of research on attentional capture by emotional faces is inconsistent. The new paradigm introduced by Hodsoll et al. (2011) and Huang et al. (2011) seems promising for further research because it allows investigating if a truly irrelevant emotional distractor affects visual search. The usefulness of this paradigm, however, depends on the careful definition of the task goal. In the study by Hodsoll et al. (2011), the target was defined in such a complex way so that parallel search seems very unlikely, whereas the target in Huang et al. (2011) was a simple feature (a black dot), which can be detected in parallel. Only if one can make an informed judgment that search was parallel (like in the additional singleton task with “meaningless objects”), one can conclude that a distraction effect was due to attentional capture. A study that combines the additional singleton task, photographic face stimuli and an easy-to-detect target attribute and that monitors eye movements is suitable to investigate whether emotional expressions really capture attention. It is striking that the study by Hunt et al. (2007) is one of the few eye tracking studies on visual search with emotional faces, and the only eye tracking study so far that also employed a task set that was independent of emotional content. Other eye tracking studies using the “discrepant face” task set were conducted by Calvo et al. (2008) and by Calvo and Nummenmaa (2008).

3 AIMS OF THE EXPERIMENTAL STUDIES

Based on the theoretical background, this doctoral thesis examined the interplay of motivation and goal-directed control in visual search with the additional singleton task. The enduring scientific debate whether attentional capture is a purely bottom-up effect or if it can be controlled by task goals (see Chapter 2.1) has completely neglected the possibility that observers' motivation to attain the task goal could determine how much effort they are willing to invest in top-down control (see Chapter 2.2). *Thus, the seemingly cognitive deficit of attentional capture might actually be a motivational one.* Whereas the possibility that motivation affects goal-directed search has almost not been considered by research, a multitude of studies using visual search tasks examined the hypothesis that stimuli with motivational content (e.g., faces with threatening or happy expressions) capture attention. Although these studies surely embarked on an important issue, most did not consider the role of top-down processes during the search tasks that they used. In line with the (cognitive) theories of visual search, this doctoral thesis argues that emotion cannot be perceived preattentively (see Chapter 2.3). *Thus, the seemingly motivational effect of emotional attentional capture is assumed to be a cognitive one because emotion was either relevant for the task set or because search was serial.* According to the active vision account (see Chapter 2.1.2), attentional selection is the result of the interaction between covert and overt processes. However, only a few studies have investigated the oculomotor components of attentional capture so far. Consequently, the experimental tasks employed in the following studies required participants to move their eyes in order to find the target and eye movements were monitored in all but one study.

To test the proposition that covert and overt attentional capture by a static, irrelevant and salient singleton is modulated by participants' motivation, three experimental studies were conducted, all of them employing different versions of the additional singleton task. It was argued that participants are normally distracted by the irrelevant singleton because the goal of distractor suppression is not important to them and they therefore do not engage in effortful top-down control (see Chapter 2.2). Following this argumentation, **Study 1 (see Chapter 4)** examined the effect of incentives for goal attainment on top-down control over covert and overt attention in the additional singleton task. It was hypothesized that monetary incentives increase the mental effort invested during search and decrease attentional and oculomotor capture by the salient singleton, in comparison to a neutral baseline. Study 1 yielded unexpected results because covert and overt capture was not observed in the baseline condition. As is outlined in the Discussion of Study 1, this raised questions of how adequate the task was to measure incentive effects on distractor suppression. Based on recent evidence of distractor learning in the additional singleton task, **Study 2 (see Chapter 5)** more closely examined the nature of practice effects in this task with a more complex design, but without eye tracking. These results led to the design of **Study 3 (see Chapter 6)**, which tested the same hypothesis as Study 1, but with a different version of the additional singleton task and with a different payoff scheme. Study 1 and Study 3 had the additional aim to investigate whether attentional capture is also driven by eye movements and whether reward affects both covert and overt attentional processes.

Finally, the role of top-down control for the emotional “counterpart” of attentional capture was examined. **Study 4 (see Chapter 7)** tested the hypothesis that attentional capture by emotional faces is not a bottom-up effect, but mediated by the top-down task and by the accompanying search processes of the participants: If the target attribute is not related to emotional content and if search can be conducted in parallel, emotional faces do not capture covert and overt attention (see Chapter 2.3). To test this hypothesis, visual search in two “emotional” variants of the additional singleton task was compared.

4 STUDY 1: THE EFFECT OF INCENTIVES ON MENTAL EFFORT AND TOP-DOWN CONTROL OVER ATTENTIONAL CAPTURE

4.1 Hypotheses

The following hypotheses were derived from the theoretical background outlined in Chapter 2.1 and Chapter 2.2:

- **Hypothesis 1.1:** Performance-contingent monetary incentives increase the mental effort invested during search in the additional singleton task, compared to a baseline condition with no incentives.
- **Hypothesis 1.2:** Performance-contingent monetary incentives increase the search performance, compared to a baseline condition with no incentives.
- **Hypothesis 1.3:** Performance-contingent monetary incentives decrease covert and overt capture by the color singleton distractor (CSD), compared to a baseline condition with no incentives.
- **Hypothesis 1.4:** The effect of high monetary incentives on mental effort, search performance and attentional capture is larger than the effect of low monetary incentives.

4.2 Method

4.2.1 Participants

Thirty-one students (24 female) of the University of Heidelberg took part in the experiment. Their age ranged from 19 to 58 years (*M* = 24). The participants received course credit and the amount of money that they won in the experiment as compensation. The monetary payment ranged from 4.93 € to 10.52 €. On average, participants won 7.80 € in the experiment (*SD* = 1.38).

4.2.2 Payoff Scheme

There were five different incentive conditions: in *high loss* trials, participants lost 10 ct of their money if they made an incorrect or slow response. In low loss trials, they lost 1 ct. In *neutral* trials, participants neither lost nor won any money. In *low reward* trials, they won 1 ct in case of a correct and fast response. In *high reward* trials, participants won 10 ct for a correct and fast response. A fast response had to be within an individual time limit. This limit was the median of the last six responses plus 50 ms. An individual and adaptive limit guarantees that the task remains demanding despite learning effects over the course of the experiment. Participants started with an "account balance" of 7.00 €. The current account balance was presented to them after each trial. Three characteristics of the payoff scheme differed from those in many other studies (see Chapter 2.2.2 and Table 1): Firstly, participants really received the amount of money that they won. However, participants were paid at least 4.00 € for participation, which was not communicated to them in advance. This method was chosen in order to guarantee a realistic setting. Due to the payoff scheme, participants were expected to "earn" on average approximately 7.00 €. This was, at the time of data collection, the standard amount of compensation for participating in an hour-long experiment at the University of Heidelberg. Secondly, participants were told that the rewards that they won during the experiment were the only form of monetary compensation for taking part in the experiment. Thirdly, participants received feedback on their performance and on the monetary amount won or lost in the trial and on the accumulated amount of reward won so far after each trial. This was added because a feedback that does not show participants how much money they currently possess could lead participants to actually doubt that any rewards can be won. Thus, the payoff scheme was similar to the study by Voss (2004), with the exception that monetary amounts, not points, were won and lost.

4.2.3 Design

Trial type and *incentive* were manipulated as within-subjects variables in this experiment. First, half of the trials contained a color singleton distractor (*CSD present* trials) and half of the trials did not (*CSD absent* trials). These conditions were combined with the five incentive conditions (see Chapter 4.2.2). Thus, there were ten different combinations of the two independent variables. In each block, the sequence of the combinations was randomly determined and repeated six times. The participants completed 36 trials of each combination during the experiment.

4.2.4 Stimuli

In every trial, six stimuli were presented on black (RGB: 0, 0, 0) background, positioned on an imaginary circle with a radius of 11.4°. The stimuli were located in the periphery of the visual field and could only be identified by making eye movements to their position. The target was a circle (diameter: 2.3°) and the distractors were diamonds (a square with a side length of 1.9° rotated by 45°). All of the objects had a line width of 0.1° and were green (RGB: 0, 78, 0), except for the CSD, which was red (RGB: 255, 0, 0). Target lines were grey (RGB: 128, 128, 128), centered inside each outline shape and had a length of 0.5° and a line width of 0.1°. Each line was rotated by 45° either to the left or to the right. Stimuli (their eccentricity, color and size) were similar to the ones used in the study by Müller et al. (2009). RGB values of the target and the CSD color were matched in their objective luminance. This is important because the CSD is supposed to be a singleton due to its unique color and not due to its unique luminance. If the CSD was less bright than the target, it would also be less salient. For each RGB value, relative luminance (Y) was computed with the following formula: $Y = 0.213 \times R + 0.715 \times G + 0.072 \times B$ (Poynton, 2003). This resulted in a relative luminance of 56 for the green color and of 54 for the red color.

4.2.5 Apparatus

Eye movements and pupil diameter were recorded with the table-mounted iView X™ Hi-Speed eye tracker by SensoMotoric Instruments, with a temporal resolution of 1250 Hz. During the experiment, participants place their head unto a chin rest, which minimizes head movements. The viewing distance between the observer and the computer screen was 60 cm. A 13 point calibration was used. The experiment was programmed in C/C++. The experiment was presented on a 23.6'' LCD monitor with a resolution of 1920 x 1080 pixels. Manual responses were recorded with highly accurate response devices (Voss, Leonhart, & Stahl, 2007).

4.2.6 Procedure

After participants had signed an informed consent and had read the instructions, the eye tracker was calibrated. 13 small squares appeared at different locations on the screen one after another and had to be fixated by the observer. In case of a successful calibration, the experimental task started. At the beginning of each trial the reward/loss cue (-10 ct, -1 ct, 0 ct, +1 ct, +10 ct) was presented for 2000 ms. This was replaced by a fixation cross (0.5° x 0.5°), which remained on screen throughout the search task. After 500 ms, the search display was presented until the observer responded. Participants' task was to press the right response key if the line inside the target (green circle) was tilted to the right (/) and the left response key if the line inside the target was tilted to the left (\). They were informed that in some trials one of the diamonds would be red and that they should ignore it. Instructions equally emphasized speed and accuracy. Immediately after their response, participants received feedback on their performance. If they gave the wrong response, the word "Fehler" (error) was presented in the middle of the screen. If their response time exceeded their individual time limit for a fast response, the words "zu langsam" (too slow) appeared. In addition, the amount of their win (+1 ct, +10 ct, in green color) or loss (-1 ct, -10 ct, in red color) in the trial and their current account balance were presented in the middle of the screen. The feedback remained on screen for 2000 ms. The next trial started after an intertrial interval of 500 ms (see Figure 1). Participants completed one practice block of 60 trials and continued with six blocks of 60 trials each. The eye tracker was calibrated at the beginning of each block, giving participants the opportunity to have a short break. One session took approximately 60 minutes.

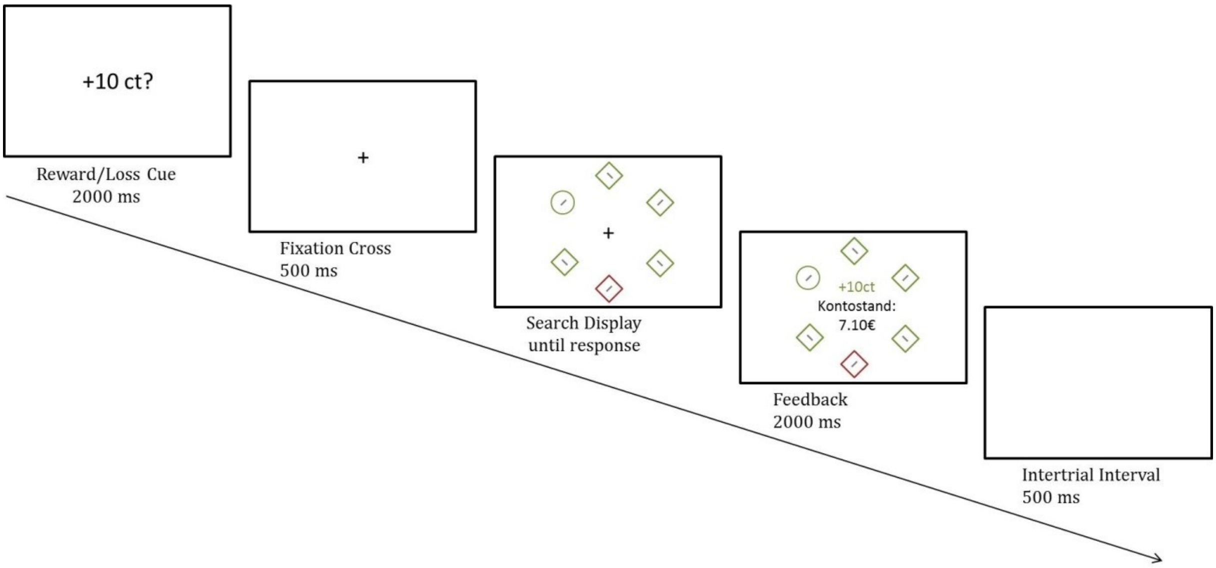

Figure 1. Schematic overview of a trial in the additional singleton task (Study 1).

4.2.7 Eye Movement Event Detection

Fixations, saccades and blinks were detected with the software BeGaze 3.4 by SensoMotoric Instruments. This software uses a high-speed event detection algorithm: each raw data sample whose velocity exceeded 30°/s was classified as a saccade and each data sample for which gaze was not recorded was classified as a blink. The remaining data samples were categorized as belonging to fixations if the minimum fixation duration exceeded 50 ms. A saccade or fixation was defined as hitting a stimulus when the saccadic endpoint or the center of the fixation fell within a circle with the radius of 2.5° around the stimulus center. As parafoveal vision extends to 5°, observers can only perceive the target line if their eye movement lands within this radius.

4.2.8 Overview of Outcome Variables

Both covert and overt attentional processes were analyzed. As the hypotheses pertained to attentional capture, which only occurs during parallel search, the *number of fixations* (1) (see Chapter 4.3.1) made during a trial served as the variable to check whether search was parallel in this task. Like in all of the other studies in the literature so far, *accuracy* (see Chapter 4.3.2) and *response latency* (see Chapter 4.3.3) were analyzed as indicators of covert attentional capture.

The indicators of overt attentional capture are less well defined: The studies that have investigated overt capture in the additional singleton task differ in the variables that they examined (Geyer et al., 2008; Theeuwes et al., 2003). Therefore, this study aimed to analyze all eye movement processes which could theoretically show signs of overt attentional capture. In analogy to response latency, a distractor could delay *saccadic latency* (see Chapter 4.3.4). The most obvious indicator of overt capture is probably the *hit rate of the first saccade on the CSD* (the percentage of trials in which the first saccade is directed to the CSD). Yet, even if the CSD is not fixated more often than another distractor, it might divert gaze away from the target and thus affect the *hit rate of the first saccade on the target* (the percentage of trials in which the first saccade is directed to the target) (see Chapters 4.3.5 - 4.3.7). Finally, the *fixation duration on the CSD* (see Chapter 4.3.8) was analyzed, too, in order to test the delayed disengagement account (Belopolsky et al., 2011; Fox et al., 2001; Fox et al., 2002). This account argues that emotional stimuli are not fixated more often than other stimuli, but delay the disengagement of attention once they are fixated. Therefore, emotional stimuli should be fixated longer than other stimuli. Although this account refers to emotional stimuli, it also points out that delayed disengagement is a general distraction effect that could be invoked by other salient stimuli, as well. The mental effort invested during the task was assumed to vary on a trial-wise basis and was therefore measured with *phasic pupil dilation* (the difference between the peak pupil dilation during the search task and the mean baseline pupil dilation during the presentation of the reward/loss cue; see Chapter 4.3.9).

4.2.9 Data Preparation

As response latencies have a positive skew, they were logarithmized for analyses. Untransformed reaction times are displayed in text, figures and tables. Trials with incorrect responses were excluded from the dataset (19.1%). To detect outlier values, z-scores were computed for logarithmic response time and saccadic latency. In accordance with Tabachnick and Fidell (2014), a z-score higher than 3.29 was considered to be an outlier. In addition, absolute cutoffs were defined to control for fast anticipatory outliers at the left tail of the distribution that would not be detected by the z-score cutoff. Response times and saccadic latencies whose z-scores exceeded 3.29 or which were faster than 200 ms (response time) or 80 ms (saccadic latency) were excluded. Saccades that were initiated earlier than 80 ms after display onset were considered as anticipatory because they must have been initiated prior to display onset. Outliers on response latency made up only 0.1% of the responses. In 7.4% of the trials saccadic latency could not be computed because the eye tracker did not register a saccade. As the target could only be correctly identified by means of eye movements, these trials probably reflect measurement errors. When the gaze position cannot be measured at the beginning of the trial, participants either blinked or the pupil position could not be detected by the image analysis software of the eye tracker. Changes in the amount of sunlight shining through the window or small head movements can change the eye video and can cause the software to lose track of the pupil during the experiment. Although the experimenter was watching the eye video on a computer, it was not always possible to correct it instantly in case of such changes[2]. Trials without eye movements and trials in which saccadic latency was higher than 0 ms, but lower than 80 ms (4.9%) were discarded. 0.4% of the trials were dismissed because z-transformed saccadic latency exceeded 3.29. In a whole, 31.9% of all responses were discarded due to these criteria. As a result, the datasets of four participants had to be excluded from the analysis because they lacked data in at least one cell of the dependent variables: The proportion of discarded trials ranged from 18% to 83% among these four participants. Although 18% is not much compared to the average proportion of discarded trials, it can lead to the dismissal of the whole dataset if the discarded trials are cumulated in one condition. The dataset of one participant was incomplete due to program failure after the first block of the experimental session and therefore had to be skipped from analysis. Thus, the datasets of 26 participants were analyzed.

4.3 Results

To test the hypotheses, two-way ANOVAs with incentive and trial type as within-subjects factors were computed. The following a-priori contrasts for the independent variable incentive were defined:

- Contrast 1 (C1) compared high loss to low loss trials.
- Contrast 2 (C2) compared high reward to low reward trials.
- Contrast 3 (C3) compared reward to loss trials.
- Contrast 4 (C4) compared valent trials (reward and loss) to neutral trials.

4.3.1 Number of Fixations

On average, participants made 1.62 fixations during a trial (95% CI [1.46, 1.78], range [0.75, 2.26]). The number of fixations was significantly lower than 2, $t(25) = -4.85$, $p < .001$, 95% CI [-0.54, -0.22], indicating parallel search.

---

[2] This issue is addressed in Study 3.

4.3.2 Accuracy

The 2 (trial type: CSD absent, CSD present) x 5 (incentive: high loss, low loss, neutral, low reward, high reward) - ANOVA on accuracy yielded a non-significant effect of trial type, $F < 1$. Contrasts 1 - 4 for incentive were not significant, $Fs(1, 25) < 3.5$, $ps > .07$. The interaction between trial type and incentive was significant for Contrast 1, $F(1, 25) = 5.15$, $p = .032$, $\eta^2_p = .17$. Contrasts 2 - 4 were not significant for the interaction, $Fs < 1$. Unexpectedly, the error rate was higher in CSD absent trials compared to CSD present trials in the high loss condition. This pattern was reversed in the low loss condition. Error rates were higher for trials with a CSD than without a CSD in the other incentive conditions (see Table 2).

Table 2. *Mean Accuracy in % (SD in Brackets), Dependent on Trial Type and Incentive (Study 1).*

	High loss	Low loss	Neutral	Low reward	High reward
CSD absent	79 (15)	82 (15)	83 (13)	81 (14)	81 (14)
CSD present	82 (14)	79 (14)	82 (14)	81 (15)	79 (12)

4.3.3 Response Latency

The 2 (trial type) x 5 (incentive) - ANOVA on response latency revealed a main effect of incentive for C2 [$F(1, 25) = 9.40$, $p = .005$, $\eta^2_p = .27$], C3 [$F(1, 25) = 5.42$, $p = .028$, $\eta^2_p = .18$] and C4 [$F(1, 25) = 44.51$, $p < .001$, $\eta^2_p = .64$]. C1 was not significant, $F(1, 25) = 3.02$, $p = .094$. Observers searched faster in high reward compared to low reward trials. In addition, response latencies were faster in the reward than in the loss conditions and were slowest in the neutral condition. There was no main effect trial type, $F < 1$. None of the contrasts was modulated by trial type, $Fs(1, 25) < 3.17$, $ps > .08$. As can be inferred from Figure 2, response latencies were only slowed by the CSD in the low loss condition. This was corroborated by a *t*-test, $t(25) = -2.35$, $p = .027$. Mean interference was 23 ms, 95% CI [- 42, -4] in this condition.

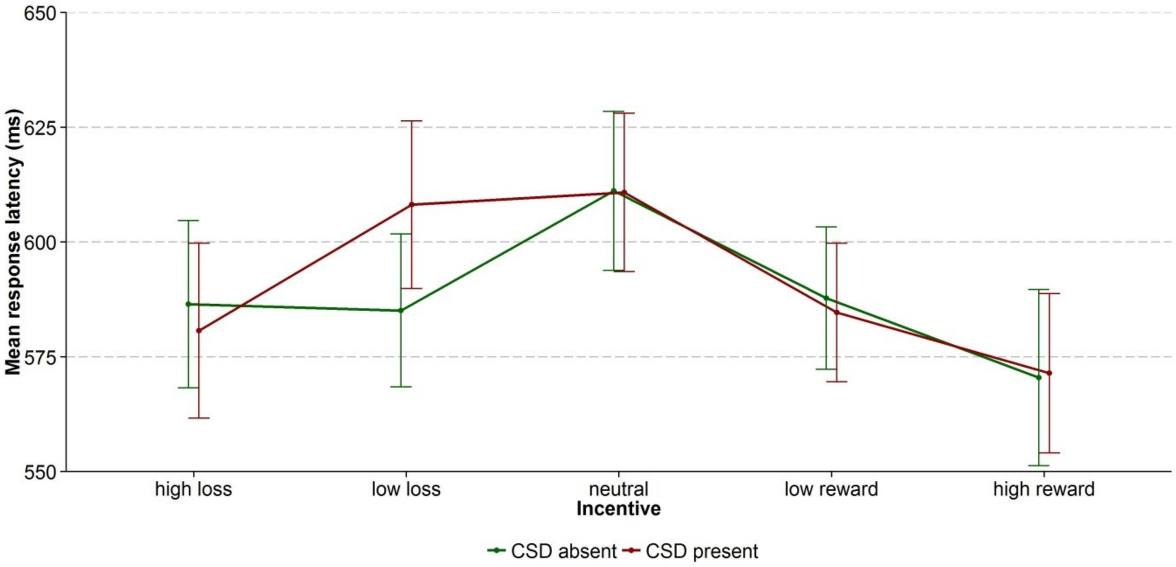

Figure 2. Mean response latencies, dependent on trial type and incentive. Error bars represent standard errors. Points are set off horizontally to enhance the visibility of the error bars (Study 1).

4.3.4 Saccadic Latency

The 2 (trial type) x 5 (incentive) - ANOVA on saccadic latency revealed a non-significant effect of trial type, $F(1, 25) = 1.65$, $p = .21$. C1 – C3 for the factor incentive were non-significant, *F*s < 1. C4 reached significance, $F(1, 25) = 5.86$, $p = .023$, $\eta^2_p=.19$, indicating that saccades to the target were executed faster in reward and loss trials compared to neutral trials. The interaction between trial type and incentive was significant for C2, $F(1, 25) = 4.96$, $p = .035$, $\eta^2_p = .17$, yet the difference between the two trial types was neither significant in the low reward condition, $t(25) = 1.30$, $p = .21$, 95% CI [-5, 23] nor in the high reward condition, $t(25) = -1.35$, $p = .19$, 95% CI [-17, 3] (see Figure 3). C1, C3 and C4 were not significant for the Trial Type x Incentive interaction, *F*s < 1.

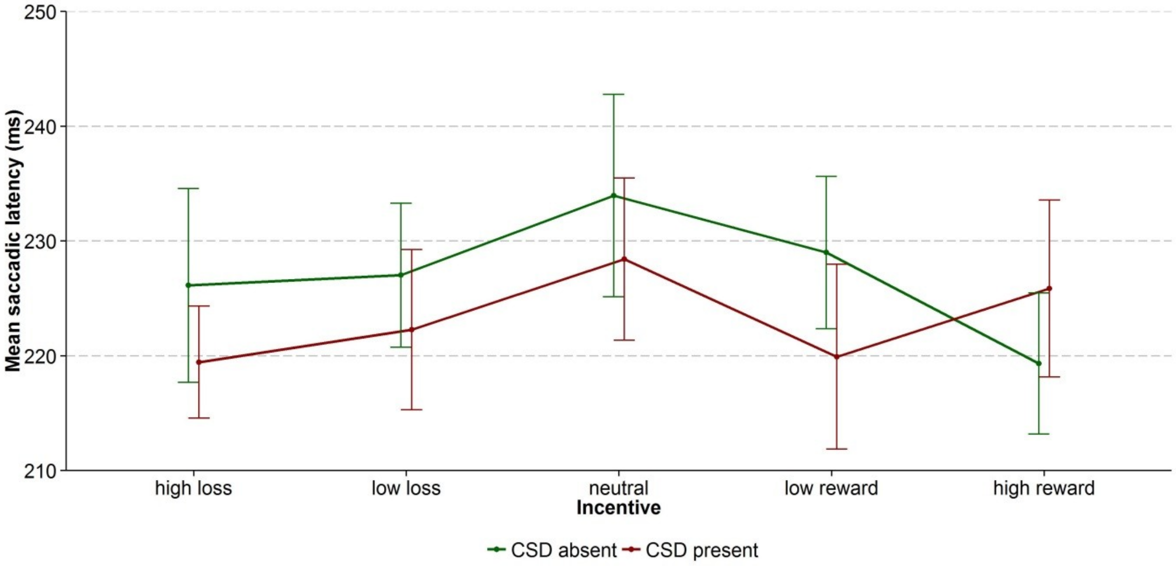

Figure 3. Mean saccadic latency, dependent on trial type and incentive. Error bars represent standard errors. Points are set off horizontally to enhance the visibility of the error bars (Study 1).

4.3.5 Location of First and Later Saccades

The random probability to hit one of the stimuli with the first saccade, if an observer selects one of the six stimulus positions as a saccade target, is approximately 17%. For mainly descriptive and illustrative purpose, the hit rates (i.e., the percentage of trials) of first and all saccades on the different stimuli (target, form distractor, CSD) for the two trial types and for the five incentive conditions are displayed in Table 3. The hit rate of the first saccade on the target refers to the percentage of trials in which the target was the landing point of the first saccade, relative to the number of trials in this condition. The hit rate of all saccades to land on a certain location refers to the percentage of trials in which this location is the landing point of any of the saccades (including the first one) during the trial, relative to the number of trials in this condition. Therefore, these values do not contain information about the number of fixations on a certain location, but express the probability that the first or any saccade fell on this location at all during a trial. The hit rate of saccades on a form distractor refers to the percentage of trials in which a saccade fell on one of the form distractors (i.e., the total hit rate on any of the form distractors divided by the number of form distractors in the display).

The hit rate of the first saccade on the target (M = 35 - 44%) lied clearly above the random hit rate in each condition. The hit rates of the first saccade on the CSD and on the form distractors, on the other hand, were well below the random probability. The hit rates for these two locations do not increase by much if all saccades are taken into account. Still, the CSD is only fixated in 2 – 6% of the trials. On the other hand, the hit rates on the targets double in size if all saccades that are executed during a trial are considered. The highest proportion of first saccades landed in between the stimuli (*M* = approximately 40%). These data suggest that the eye movements were preferentially guided to the target and seemed to avoid distractor positions (see Table 3).

Table 3. *The Hit Rates of the First Saccade and of All Saccades on Different Stimuli (SD in Brackets), Dependent on Trial Type and Incentive (Study 1).*

		First saccade			All saccades		
		TA	FD	CSD	TA	FD	CSD
High loss	CSD absent	38 (17)	4 (2)	---	77 (19)	4 (3)	---
	CSD present	35 (14)	5 (3)	5 (6)	77 (20)	6 (3)	5 (6)
Low loss	CSD Absent	36 (18)	4 (3)	---	74 (21)	4 (3)	---
	CSD present	35 (14)	5 (3)	3 (3)	78 (16)	6 (3)	4 (4)
Neutral	CSD absent	44 (19)	4 (2)	---	83 (15)	4 (2)	---
	CSD present	34 (13)	4 (3)	2 (3)	77 (19)	5 (3)	3 (3)
Low reward	CSD absent	37 (17)	4 (3)	---	77 (20)	5 (3)	---
	CSD present	37 (17)	5 (3)	4 (4)	77 (18)	5 (3)	4 (4)
High reward	CSD absent	38 (21)	4 (2)	---	80 (17)	4 (3)	---
	CSD present	38 (20)	5 (4)	2 (2)	78 (17)	6 (4)	2 (3)

Note. TA = target, FD = form distractor, CSD = color singleton distractor.

4.3.6 Comparison of the Hit Rate of the First Saccade on the CSD to the Hit Rate of the First Saccade on the Form Distractors

Although the hit rates of the first saccade on the form distractor and on the CSD were relatively low (see Table 3), it was tested whether they differed in CSD present trials. An ANOVA with saccade location (form distractor, CSD) and incentive as within-subjects factors revealed a large effect for saccade location, $F(1, 25) = 12.69$, $p < .002$, $\eta^2_p = .34$, with a higher proportion of first saccades landing on one of the form distractors than on the CSD. Furthermore, the interaction between saccade location and incentive was significant for C2, $F(1, 25) = 6.54$, $p = .017$, $\eta^2_p = .21$. The remaining effects did not reach significance, *F*s < 3.85, *p*s > .06. As illustrated in Figure 4, the difference between the CSD and the form distractor hit rates was higher in the high reward condition than in the low reward condition. These results show that observers effectively ignored the CSD with their eyes, especially in high reward trials.

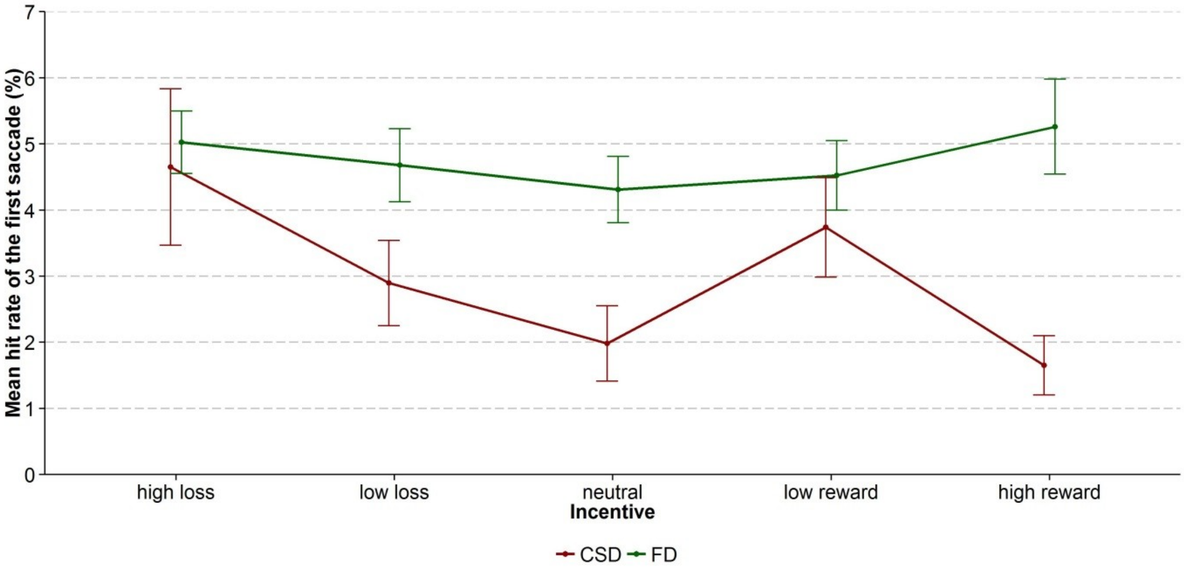

Figure 4. Mean hit rate of the first saccade on a form distractor (FD) and on the CSD, dependent on incentive. Error bars represent standard errors. Points are set off horizontally to enhance the visibility of the error bars (Study 1).

4.3.7 Hit Rate of the First Saccade on the Target

As the hit rate of the first saccade on the CSD was significantly lower than the hit rate of the first saccade on the form distractors, the CSD did not directly capture overt attention. However, the CSD could have captured attention indirectly by reducing the target hit rate. The 2 (trial type) x 5 (incentive) - ANOVA on the target hit rate of the first saccade revealed a marginally significant main effect of trial type, $F(1, 25) = 4.06$, $p = .055$, $\eta^2_p = .14$. The main effect of incentive was non-significant for C1 - C4, *F*s < 1.53, *p*s > .22. The interaction between trial type and incentive was significant for C4, $F(1, 25) = 9.74$, $p = .005$, $\eta^2_p = .28$. The remaining contrasts of the interaction were not significant, Fs < 1. As can be inferred from Figure 5, the target hit rate differs between the trial types only in the neutral condition, where the hit rate is much higher for CSD absent trials than for CSD present trials. In contrast to response latency and saccadic latency, the target hit rate is not improved in incentive conditions in comparison to the neutral baseline. The hit rate in CSD present trials is on a similar level in all of the five incentive conditions. Therefore, the increase of attentional capture in the neutral condition compared to the valent conditions is atypical: it is not due to a performance decrease in CSD present trials, but due to a performance increase in CSD absent trials.

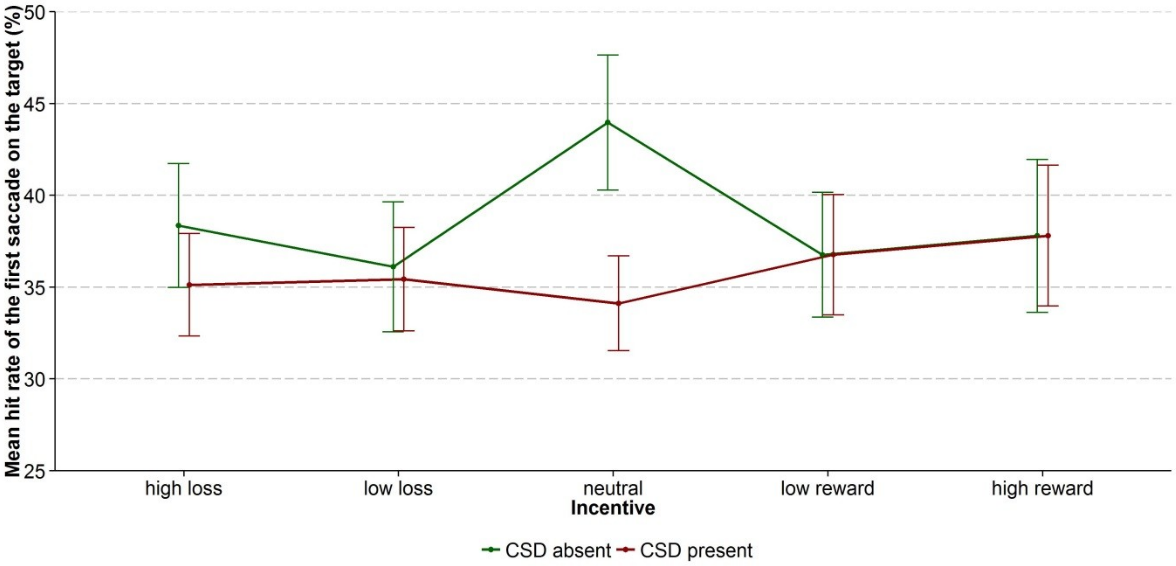

Figure 5. Mean hit rate of the first saccade on the target, dependent on trial type and incentive. Error bars represent standard errors. Points are set off horizontally to enhance the visibility of the error bars. (Study 1).

4.3.8 Fixation Duration on the CSD

Due to the low rate of fixations on the CSD (see Chapter 4.3.5), fixation duration was not analyzed as a dependent variable.

4.3.9 Phasic Pupil Dilation

Phasic pupil dilation was computed as the difference between the peak pupil diameter during the time between display onset and the manual response and the mean baseline pupil dilation during the presentation of the reward/loss cue at the start of the trial. Thus, phasic pupil dilation indicates how much the pupil dilates during the search task. The 2 (trial type) x 5 (incentive) - ANOVA on phasic pupil dilation yielded a non-significant effect of trial type, $F < 1$. The main effect of incentive was significant for C1 [$F(1, 25) = 45.32$, $p < .001$, $\eta^2_p = .64$], C2 [$F(1, 25) = 13.17$, $p = .001$, $\eta^2_p = .35$], C3 [$F(1, 25) = 5.97$, $p = .022$, $\eta^2_p = .19$], and C4 [$F(1, 25) = 60.83$, $p < .001$, $\eta^2_p = .71$]. Pupil dilation was increased in high loss compared to low loss trials (C1), in high reward compared to low reward trials (C2) and in valent compared to neutral trials (C4). Furthermore, the pupil widened more in the reward conditions than in the loss conditions (C3). None of the contrasts was modulated by trial type, *F*s < 1.87, *p*s > .18 (see Figure 6).

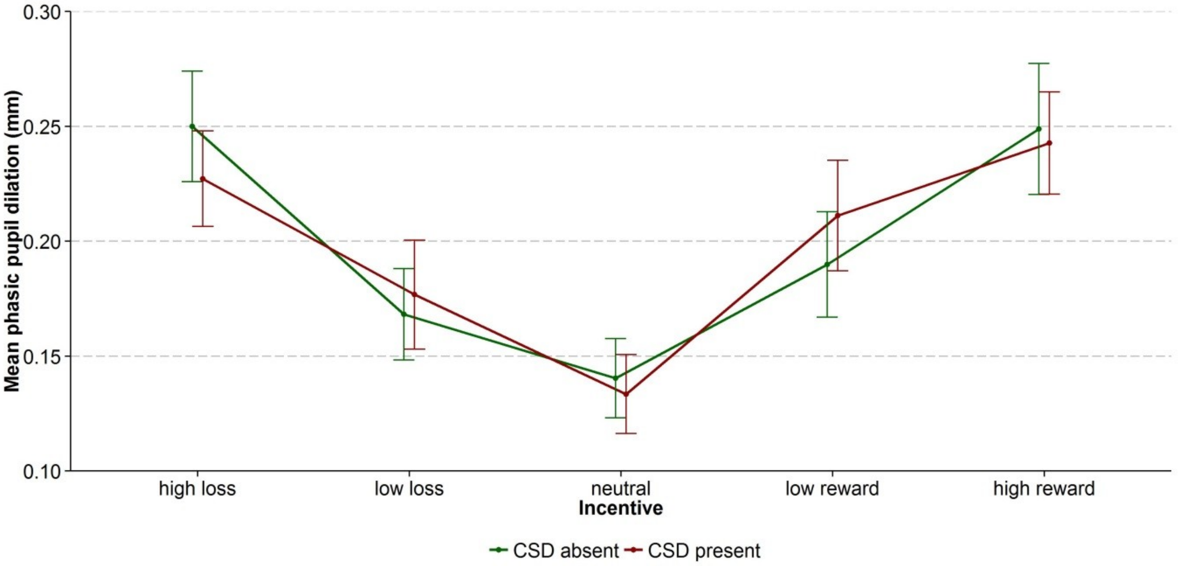

*Figure 6***.** Mean phasic pupil dilation, dependent on trial type and incentive. Error bars represent standard errors. Points are set off horizontally to enhance the visibility of the error bars (Study 1).

Contrasts 1 - 4 were significant for the main effect of incentive both for response latency and pupil dilation. Difference scores between the neutral condition and the high reward condition were computed for both dependent variables and entered into a bivariate correlation. The correlation between the two variables was not significant, *r*(26) = .07, *p* = .75. Consequently, the decrease in response latency in high reward trials did not correlate with an increase in pupil dilation.

4.4 Discussion

It was hypothesized that performance-contingent monetary incentives increase the mental effort invested during search in the additional singleton task (Hypothesis 1.1) and increase the search performance (Hypothesis 1.2). Importantly, it was expected that incentives decrease the amount of covert and overt attentional capture by the CSD (Hypothesis 1.3). Furthermore, Hypothesis 1.4 assumed that high monetary rewards have larger effects on mental effort, search performance and attentional capture than low monetary rewards. To arrive at any conclusions about top-down modulability of attentional capture, one has to make sure that the visual search task was adequate to measure attentional capture. This is only the case if the target can be found in parallel search. Instead of varying set size, which is the standard method to compute the efficiency of search, the number of fixations executed during search was taken as an indicator. If participants needed, on average, two or less fixations to find the target, it can be assumed that search was parallel (see Chapter 2.1.2). In this study, the mean number of fixations was significantly lower than two. Therefore, any distractor interference in overt or covert indicators of attention is due to attentional capture. The next two chapters (4.4.1 and 4.4.2) discuss the results of Study 1 and relate them to the hypotheses and to previous results in the literature. The third chapter (4.4.3) summarizes the results with regard to the interaction between covert and overt attention. The fourth chapter (4.4.4) addresses some limitations of the study and the last chapter (4.4.5) presents the conclusions that were drawn from the results and that motivated Study 2.

4.4.1 Monetary Incentives Increase Mental Effort in Parallel Search and Decrease Response Latencies and Saccadic Latencies

In line with Hypothesis 1.1, incentives were found to increase mental effort, as indicated by an increase in phasic pupil dilation both in reward and loss conditions. If participants were cued with the possibility of winning or with the danger of losing money, they invested more effort during search. Furthermore, in line with Hypothesis 1.4, mental effort was increased in high incentive conditions compared to low incentive conditions. These results are consistent with other studies showing that incentives enhance mental effort instantly (Bijleveld et al., 2009; Chiew & Braver, 2013).

Furthermore, Hypothesis 1.2 was confirmed because incentives decreased response latencies and saccadic latencies. Observers detected the target faster in high loss and in high reward trials than in low loss and low reward trials. This facilitation effect was not paralleled by a decrease in accuracy or in the hit rate of the first saccade on the target and thus reflects a performance benefit caused by incentives. Accordingly, the figures for response latency, saccadic latency and pupil dilation (see Figure 2, Figure 3, and Figure 6) show a quadratic trend of the factor incentive. Based on the theoretical background, the obvious explanation for the facilitation effect of incentives on response latency and saccadic latency is that the incentives motivated participants to attain the task goal (i.e., finding the target) and thereby led to the investment of more effort for effective top-down control. This interpretation is supported by studies showing that reward enhances cognitive control in response conflict tasks (Chiew & Braver, 2013; Fröber & Dreisbach, 2014; Locke & Braver, 2008; Savine et al., 2010; Savine & Braver, 2010; Veling & Aarts, 2010; see Chapter 2.2.2).

Response latencies were also faster in reward trials compared to loss trials, indicating that the expectancy to win is more motivating than the expectancy to lose. The same pattern was found in the hit rate of the first saccade on the CSD because participants made fewer saccades towards the CSD if they could win than if they could lose. This suggests that the finding that “losses loom larger than gains” (Kahneman & Tversky, 1984, p. 346) does not apply to visual search, or to the small amounts of money that were won and lost in this experiment. The only other study on visual attention that employed a payoff scheme with rewards and losses did not observe a difference between them (Engelmann & Pessoa, 2007).

The decrease in response latencies from the neutral to the high reward condition, however, did not correlate with the increase in mental effort from the neutral to the high reward condition. Either participants recruited more effort than they actually needed for effective top-down control or phasic pupil dilation measures not only mental effort. This question is further discussed in Chapter 8. The next chapter (4.4.2) discusses the results of Study 1 with regard to Hypothesis 1.3, which assumed that incentives decrease the amount of covert and overt attentional capture.

4.4.2 The Surprising Absence of Attentional Capture

The most surprising finding of this study was that the CSD did not delay response latencies in the neutral condition. It was expected that trial type and incentive interact for the comparison of the neutral condition and the incentive conditions (Contrast 4): Attentional capture was expected to be lower in reward and loss trials compared to neutral trials, in which participants did not receive any incentive to suppress the distractor. In contrast to the hypothesis, the results showed that the CSD only interfered with response latency in the low loss condition and not in the neutral condition, in which pupil dilation was lowest and in which attentional capture was expected to be highest. The CSD slowed response time in the low loss condition by 23 ms, which is comparable to the attentional capture effect reported by Theeuwes (1992). Covert distractor interference in the low loss condition was not paralleled by overt interference.

The oculomotor data even demonstrated that the CSD was intentionally avoided by participants' gaze. Observers directed their first saccade to the CSD only in 2 - 5% of the trials, which lied well below the random hit rate of 17% and which was even significantly lower than the hit rate on the form distractor. Therefore, the target clearly had the greatest impact on the first orientation of overt attention. The hit rate on the CSD is similar in size to the one reported by Theeuwes et al. (2003, Exp. 2), but, in contrast to that study, covert attentional capture was absent in the present study.

An explanation of these results could be that the CSD did not capture attention in the neutral condition because it was not salient enough and that capture in the low reward condition was an artifact. Theeuwes (1992) assumed that the CSD does not capture attention if its bottom-up activation is lower than that of the target. Several aspects of the results of this study, however, do not support this explanation. Firstly, the oculomotor data show signs of overt attentional capture. The CSD did not directly capture overt attention, but observers made significantly less often a first saccade towards the target in the neutral condition when the CSD was present. Thus, the distractor indirectly captured overt attention in this condition. Secondly, RGB values of the target and the CSD color were matched in their objective luminance in order to guarantee that the CSD was more salient because of its unique color. In order to verify that the CSD was not only objectively more salient, but also subjectively perceived as more salient than the target, response latencies in the practice block (30 trials of each trial type) were analyzed. This revealed a significant capture effect of the CSD, $t(25) = -2.46$, $p = .02$. Responses were slowed by 23 ms in CSD present trials (95% CI [2, 44]). The size of the interference effect is comparable to the one reported by Theeuwes (1992) and to the one found in the low loss condition in this study. This shows that the distractor was sufficiently salient to evoke attentional capture and indicates that attentional capture was alleviated very fast during the test trials.

The absence of attentional capture in the test trials is especially striking because van Zoest et al. (2004) and Siebold et al. (2011) observed that attentional capture is dependent on response speed: the slower the response, the smaller attentional capture. Whereas early attentional deployment immediately after stimulus onset should be completely driven by bottom-up control, responses that are initiated later in time are controlled by top-down strategy. In this study, mean response times in the practice block (743 ms in CSD absent trials, 766 ms in CSD present trials) were much slower than response times in the test trials (588 ms in CSD absent trials, 591 ms in CSD present trials). Mean response latencies were also below 600 ms in the study by Theeuwes (1992). Without considering reward, one therefore would have expected attentional capture to be increased in test trials, compared to the practice trials.

The search mode account (Bacon & Egeth, 1994; see Chapter 2.1.4) does not provide a coherent explanation for the lack of attentional capture in the neutral condition and for the presence of attentional capture in the low loss condition. If participants had been motivated to adopt feature search mode after the practice block in order to effectively speed up their response to the target and ignore the distractor, this search mode should have been activated in each trial. The search mode account does not make any propositions about the motivational states of the observer and therefore does not provide any explanation for why participants should switch to singleton search mode only in the low loss condition.

To conclude, the most reasonable explanation for the lack of attentional capture is that participants effectively suppressed the CSD in almost every trial and that the occurrence of attentional capture in the low loss condition is a random finding. This supports the dimension weighting account (see Chapter 2.1.5) and challenges the stimulus-driven capture account (see Chapter 2.1.3).

4.4.3 The Interaction of Covert and Overt Indicators of Attention

It was an additional aim of this study to investigate whether attentional capture is mainly a covert phenomenon, as suggested by the results of Theeuwes et al. (2003) and Geyer et al. (2008), or whether overt processes are also at play. Unexpectedly, covert attentional capture was only observed in the low loss condition and this effect was not associated with any overt interference effect. On the other hand, this study observed an atypical overt attentional capture effect in the neutral condition: The hit rate of the first saccade on the target was higher in CSD absent trials than in CSD present trials. This effect was termed atypical because the hit rate in CSD present and CSD absent trials was the same in the remaining conditions. Therefore, the distractor effect arises – in comparison to the other conditions – because of a performance increase in CSD absent trials, not because of a performance decrease in CSD present trials. In neutral trials, saccadic latencies and response latencies were slower in comparison to the other conditions. Studies showed that slower saccadic responses are less affected by distractor salience (Theeuwes et al., 2003; van Zoest et al., 2004). Based on these results, one would expect the hit rate of the target to be increased both in CSD absent and CSD present trials because saccadic latency was significantly slower in both trial types compared to the valent conditions. However, participants only increased their target hit rate in CSD present trials. This suggests that they were indirectly distracted by the CSD. This covert CSD effect was not observed in response latencies. The fact that the covert and the overt interference effect occurred in different trial types suggests that covert and overt processes of attention are, to a certain extent, independent of another, even in parallel search. This issue certainly requires further research. Furthermore, the interpretation of the distraction effect in the low loss condition is complicated because it was not expected by the hypotheses. The attentional capture effect in the target hit rate is more consistent with the theoretical reasoning: Participants put less effort into top-down control in the neutral condition, as indicated by a decrease in pupil dilation, and are therefore more distracted by the CSD.

The incentive effects on covert and overt indicators of attentional orienting were consistent: Manual and saccadic response times were decreased in valent trials compared to neutral trials. The fact that saccadic response times were affected by incentives is especially interesting because it demonstrates that top-down control also modulates the early phase of visual orienting. Previous research on reward effects (Chiew & Braver, 2013; Fröber & Dreisbach, 2014; Locke & Braver, 2008; Savine et al., 2010; Savine & Braver, 2010; Veling &

Aarts, 2010) has investigated response latencies, which, in general, do not allow inferences about the time course of reward effects.

4.4.4 Limitations

As has been described in Chapter 4.2.9, 7.4% of the trials had to be excluded from analysis because they did not contain eye movement data. In order to test whether this loss of trials caused a bias in the data[3], all of the trials in which a correct manual response was recorded were analyzed. An ANOVA with trial type and incentive as repeated measures variables yielded the same pattern like the analysis that excluded trials without recorded eye movements. The effect of trial type was clearly non-significant, $F < 1$. Thus, the exclusion of these trials did not lead to a bias in the response time data.

Nevertheless, the relatively high number of trials that had to be discarded because the eye tracker did not register any saccade requires further scrutiny. Obviously, the calibration at the start of each block was not sufficient to ensure a valid recording throughout the block. In some trials, the image analysis software was not able to determine the pupil position, most likely because of head movements or light changes. Nevertheless, the trial procedure continued, irrespective of the quality of the eye image. The integration of a "fixation check procedure" into the experimental software might be a solution to this problem. Such a procedure ensures that the trial only starts if the eye tracking software recognizes a valid and stable fixation on the fixation cross in the middle of the screen at the beginning of the trial. This procedure was therefore implemented in the following eye tracking studies (Study 3 and Study 4).

The accuracy was markedly lower in Study 1 (79 - 83%) than in previous studies using the additional singleton task, most of which reported accuracies between 90% and 95% (Bacon & Egeth, 1994; Geyer et al., 2008; Leber & Egeth, 2006; Müller et al., 2009; Theeuwes, 1991, 1992, 2004; Theeuwes et al., 2003; Wienrich & Janczyk, 2011; Zehetleitner et al., 2012). The main reason for the relatively low accuracy in this study is probably that slow, but correct responses were punished with losses in loss trials and did not result in gains in reward trials. The time limit was the median of the last six response times plus an additional 50 ms. This means that an observer could not win or prevent losses in each trial because she could not constantly improve her response time and she could not exceed some baseline level of correct response latency. It is important to know that participants were not informed about the fact that the limit was adaptive and that it was impossible to win in every trial. Therefore, participants might have tried to constantly decrease their response time at the cost of accuracy.

---

[3] By chance, the trials that were discarded might have been exactly those trials in which attentional capture occurred.

It might seem surprising that participants did not fixate the target in each trial in which they gave a correct manual response, although it was assumed that foveal fixation of the target was necessary to identify it. This result suggests that participants either guessed in some trials in order to respond faster, or that foveal fixation was not needed to identify the target. In order to examine this further, accuracies in trials with a target saccade and in trials without any target saccade were compared and found to be significantly higher in the former (M = 83% vs. M = 79%), $t(25) = -2.77$, $p = .01$, 95% CI [-.07, -.01]. Consequently, although fixation of the target increased accuracy, it did not prevent participants from making quite a lot of errors (M = 17%). This finding might seem odd at first glance, but a saccade towards the target does not automatically imply that the target item was processed deeply during the fixation. Participants probably put themselves under a strict time pressure that did not allow them to process the target sufficiently in every trial. This argumentation was confirmed by a comparison of fixation durations on the target between error trials and correct trials. On average, the target fixations were significantly briefer if the manual response was erroneous (114 ms) than if it was correct (151 ms), $t(25) = 7.32$, $p < .001$, 95% CI [27, 47].

Unfortunately, the studies which most clearly resemble this study in display and design (Geyer et al., 2008; Theeuwes et al., 2003) do not report data on the hit rate of saccades (first and later) on the target nor on the fixation duration on the target. Therefore, it is not possible to determine the minimum time needed to process the target line. In his review on eye movements in different disciplines of psychology, Rayner (2009) argued that the average fixation duration during visual search is 180 - 275 ms, but he did not further differentiate between target and distractor fixation durations and between the multitudes of visual search tasks, or, a differentiation that is important for this study, between parallel and serial search. He also did not name the studies from which he computed these mean values. Nevertheless, the range named by Rayner supports the view that a fixation of 114 ms might be too short to perceive the target item correctly. This result suggests that the payoff scheme of this study was not optimal because errors were punished in loss trials, but not in reward or neutral trials. Consequently, participants might have been motivated to accept a relatively high error rate at the benefit of decreased response latencies.

4.4.5 Conclusion

This study aimed to test whether attentional capture is affected by participants' motivation to invest the mental effort required for goal-directed attentional control. The results provided a mixed picture. The most surprising finding was that covert attentional capture was only observed in low loss trials. Since the baseline covert attentional capture effect was not found, this study could not provide a test of its modulability in the reward conditions. The CSD indirectly captured overt attention in neutral trials because target saccades were less likely in CSD present trials. On the other hand, incentives decreased response and saccadic latencies and increased pupil dilation, which is a strong indication for a facilitation effect of incentives on top-down control.

The present results are similar to that of Müller et al. (2009, Exp. 2), who did not report attentional capture with a CSD probability of 50% if participants had sufficient distractor practice in the first block and if the incentive to employ distractor suppression was high. The design of Study 1 was guided by the assumption that the manipulation of monetary incentives would not be as strong as the incentive and practice effects employed by Müller et al., but the results suggest that they were indeed. The absence of attentional capture in Study 1 points to the possibility that distractor practice was effective in reducing distractor interference in such a short period of time that a modulation of attentional capture could not be detected. The present results suggest that attentional capture is strongly modulated by practice effects, even if the display allows for the use of singleton search mode (Müller et al., 2009). In the study by Müller et al. (2009), distractor practice was manipulated by varying the CSD probability during the first block of trials: The CSD either appeared in none or in 100% of the trials. The incentive to suppress the distractor was manipulated by varying the CSD probability during the later trials: a CSD occurred either in 20%, 50% or 80% of the remaining trials. The group with an initial distractor practice in 100% of the trials and that encountered the distractor quite often (50% and 80%) did not show attentional capture. Study 1 did not manipulate distractor practice, but employed strong incentives for distractor suppression. Pupil dilation, response latency and saccadic latency revealed that those were effective in both trial types. Thus, a possible explanation for the absence of attentional capture could be that initial, moderate distractor practice (50% CSD trials), which is followed by monetary incentives, is sufficient to enable effective distractor suppression. It is important to note here that Müller et al. did not use a practice group with 50% distractor probability. Thus, this amount of initial practice might be sufficient for participants to suppress capture later on. As a conclusion, it was necessary to investigate practice effects in the additional singleton task more closely without manipulating monetary incentives in order to determine an adequate design for a study examining the modulation of attentional capture by incentives.

A recent study (Vatterott & Vecera, 2012) revealed that practice effects are much more profound than even the dimension weighting account (Müller et al., 2009; Zehetleitner et al., 2012) would suggest. In this study, participants looked for a constant target among heterogeneous distractors. Accordingly, participants had to use feature search mode (cf. Bacon &

Egeth, 1994) to find the target. In contrast to previous studies with the additional singleton task, the CSD changed its color after each of the four experimental blocks. The authors aimed to examine more closely how feature search mode works: Does it entail an abstract attentional set that facilitates the suppression of any irrelevant distractor, or is it specific with regard to the distractor color? The results revealed that the CSD only prolonged reaction times in the first half of each block (immediately after the CSD changed its color), but not in the second half. After an initial stage of attentional capture in the first half of the block, observers quickly utilized their distractor experience to suppress the CSD in the second half. Secondly, attentional capture did not decrease over the course of the experiment, but was as high in block 4 as in block 1. The authors regarded this as evidence that feature search mode works not by the establishment of an abstract attentional set (i.e., “Avoid all distractors.”) that can easily be transferred to new distractor features, but is tied to specific distractor features (i.e., “Avoid the red distractor”). This study provided the first - published - block-wise analysis of attentional capture in the additional singleton task. The importance of the results is underlined by the number of trials that the authors used. Each block consisted of only 48 trials. Thus, observers had already learned to effectively suppress the distractor after 24 trials.

This fast learning effect supports the interpretation of the results of Study 1: Attentional capture did not occur in the test blocks of Study 1 because participants had learned to suppress the CSD during the 60 trials of the practice block. The search tasks used in Study 1 and in the study by Vatterott and Vecera (2012), however, differ in one decisive aspect: Whereas participants in the former were able to use singleton search mode (because the distractors were homogeneous), they probably employed feature search mode in the latter (because the distractors were heterogeneous). Vatterott and Vecera (2012) argued that the observed fast learning effects are an essential component of feature search mode, and the authors did not propose that the same learning effects might occur in singleton search mode. According to the dimension weighting account (Müller et al., 2009; Zehetleitner et al., 2012), however, distractor practice is an essential component of any search strategy. The results of Study 1 support this account because participants were not induced to use feature search mode and still were very fast at suppressing the CSD. On the other hand, search mode was not manipulated in Study 1. Therefore, Study 2 tested the hypothesis that the pattern of results observed by Vatterott and Vecera (2012) is also found with the display and search task used in Study 1. Another aim of Study 2 was to verify that search was parallel in the variant of the additional singleton task that was used in Study 1, by varying set size.

Two recent studies reported correlations between observers’ visual working memory capacity and their amount of attentional capture: the higher the working memory capacity, the lower the attentional capture. The first study by Fukuda and Vogel (2009) employed a spatial cueing task: Participants had to indicate the orientation of a target item. In some trials, a salient distractor briefly flashed at a different position. Event-related potentials indicated that participants with low working memory capacity allocated their attention to the same extent to the target and to the distractor location. Observers with high working memory capacity, however, allocated their attention preferably to the target position. A study by Anderson and Yantis (2012) employed an additional singleton task very similar to the one used in Study 1 and reported a correlation between attentional capture and working memory capacity. Studying individual differences is a relatively new approach in research on visual selection since it has always focused on general principles of human perception. Considering the conflicting evidence in favor of bottom-up vs. top-down control and the surprising results of Study 1, measuring visual working memory capacity might be a promising way to uncover a modulating variable of attentional capture.

5 STUDY 2: THE EFFECT OF DISTRACTOR PRACTICE ON ATTENTIONAL CAPTURE IN SINGLETON SEARCH MODE

5.1 Hypothesis

Study 2 was implemented to test the following hypothesis that was derived from the unexpected results of Study 1:

- **Hypothesis 2.1:** Attentional capture in the additional singleton task with singleton search mode (i.e., with homogeneous distractors) is modulated by practice with the CSD feature. When the CSD color changes with each block, attentional capture is only observable in the first half of each block.

5.2 Method

Study 2 consisted of two independent tasks. Participants first completed a change detection task to measure their visual working memory capacity. This was followed by an additional singleton task.

5.2.1 Participants

Twenty-four students (17 female) of the University of Heidelberg took part in the experiment. Their age ranged from 22 to 52 years (*M* = 27). They were either paid 5 € or received course credit as compensation for taking part in the study.

5.2.2 Design

In the additional singleton task, two display types were compared. The *feature display* was the same as in the study by Vatterott and Vecera (2012): Participants searched for a green circle among heterogeneous distractors (one square, one triangle, and several diamonds). The *singleton display* was the same as in Study 1: Participants looked for a green circle among homogeneous distractors (diamonds) (see Figure 7). Whereas the feature display was assumed to induce feature search mode, the singleton display was expected to induce singleton search mode. Participants were randomly assigned to one of the two *display groups*. In addition, *display size* (six, nine or 12 items) was varied to test whether search was parallel or serial. In half of the trials, a CSD appeared (*trial type: CSD absent, CSD present*). This resulted in six different combinations of the two within-subjects variables. Each combination was repeated 36 times in the experiment (test blocks). The experiment consisted of one practice block (without CSD and with different set sizes) and four test blocks. Each of the test blocks contained a differently colored CSD. The sequence of CSD colors was counterbalanced across participants and the blocking of CSD color was not communicated explicitly to participants. The 36 repetitions of the Trial Type x Set Size combinations were dispersed randomly across the four blocks. For analyses, each block was split into two *sub-blocks* to examine whether distractor interference was only present early on after each color change. Thus, the design consisted of five independent variables: display group, set size, trial type, block and sub-block.

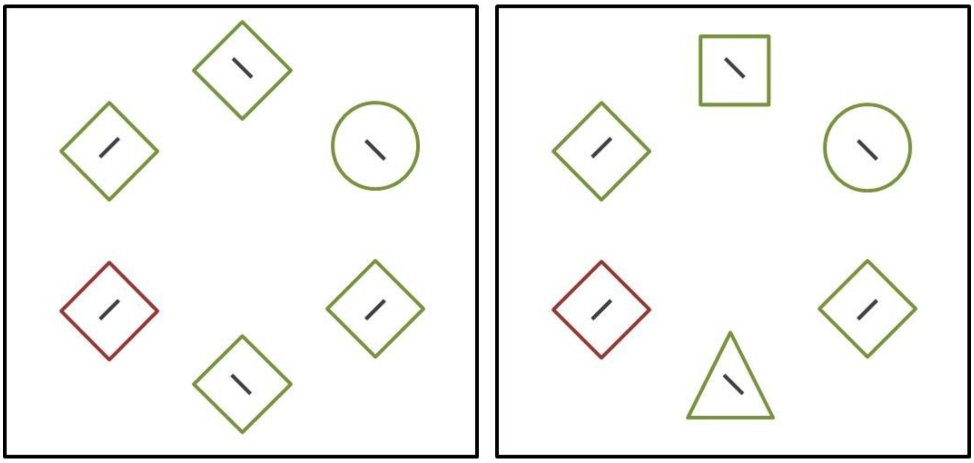

Figure 7. Schematic illustration of the singleton search display (left) and the feature search display (right) with a red CSD and display size six (Study 2).

5.2.3 Stimuli

Change Detection Task

The task was very similar to that used by Vogel, Woodman, and Luck (2001). In each trial either four, six or eight colored squares (side length: 0.6°) were presented at random positions in an imaginary square (side length: 6.6°) on grey background. The color of each square was chosen randomly from one of seven colors (red, dark blue, green, yellow, black, violet, light blue).

Additional Singleton Task

In each trial six, nine, or 12 stimuli were presented on black (RGB: 0, 0, 0) background, positioned on an imaginary circle with a radius of 11.4° (i.e., the same radius as in Study 1). The middle of the screen was marked by a fixation cross of 0.5° width and height. The following objects were used: a circle (diameter: 2.3°), a square (side length: 1.9°), a diamond (a square rotated by 45°) and a triangle (side length: 2.4°). The circle and the diamond were identical to the stimuli used in Study 1. All objects had a line width of 0.1°. Stimuli were green (RGB: 0, 78, 0) except for the CSD, which could be red (RGB: 255, 0, 0), yellow (RGB: 255, 255, 0), pink (RGB: 255, 0, 255) or orange (RGB: 255, 150, 0). These RGB values were taken from the study by Vatterott and Vecera (2012). The target lines were identical to the ones used in Study 1.

5.2.4 Apparatus

Like in Study 1, responses were recorded with self-made response devices (Voss et al., 2007). Both tasks were programmed in C/C++ and run on 23.6'' LCD monitors with a resolution of 1920 x 1080 pixels. Eye movements were not recorded in this study in order to enable fast data collection.

5.2.5 Procedure

Change Detection Task

After having signed an informed consent, both groups of participants started with the change detection task. They began with eight practice trials and continued with three blocks of 40 trials each. At the beginning of each trial a central fixation cross appeared for 200 ms. Then four, six or eight stimuli were presented on screen for 150 ms. After a retention interval of 900 ms, the same number of stimuli was presented again. In half of the trials, one of the items in the display changed its color. Participants had to press the right key if they thought that one of the items had changed its color or to press the left key if they thought that no change had occurred. If participants made an error, the word “Fehler” (error) was presented in the middle of the screen. The intertrial interval was 1000 ms (see Figure 8).

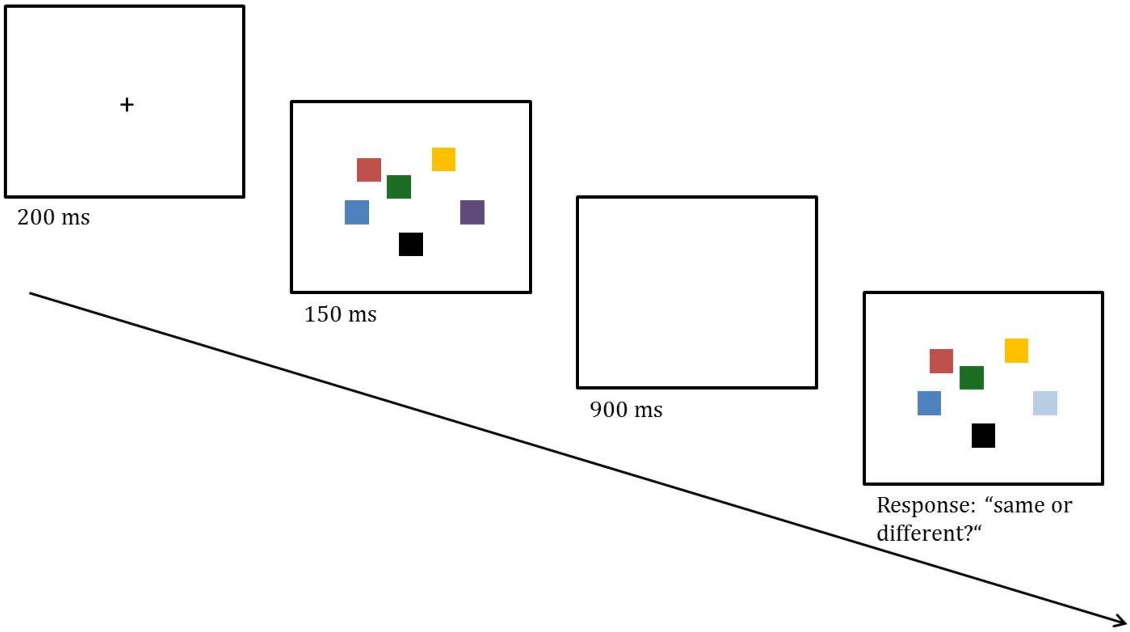

Figure 8. Schematic illustration of the trial procedure with six stimuli in the change detection task (Study 2).

Additional Singleton Task

After a short break, participants continued with the additional singleton task. Every trial began with a central fixation cross, which disappeared after 500 ms. Then, six, nine or 12 stimuli were presented and remained on screen until the observer responded. Participants' task was to press the right response key if the line inside the target (green circle) was tilted to the right (/) and the left response key if the line inside the target was tilted to the left (\). They were informed that in some trials one of the distractors would be of a different color than green and that they should ignore it. Instructions equally emphasized speed and accuracy. In case of an incorrect response the word "Fehler" (error) was presented for 500 ms. The intertrial interval was 1000 ms. Like in the study by Vatterott and Vecera (2012), both display groups first completed a practice block of 60 trials, which did not comprise a CSD. They then continued with four blocks of 54 trials each. After each block, the color of the CSD changed. After the second block, participants were allowed to take a short break. The whole experimental session took approximately 25 minutes (see Figure 9).

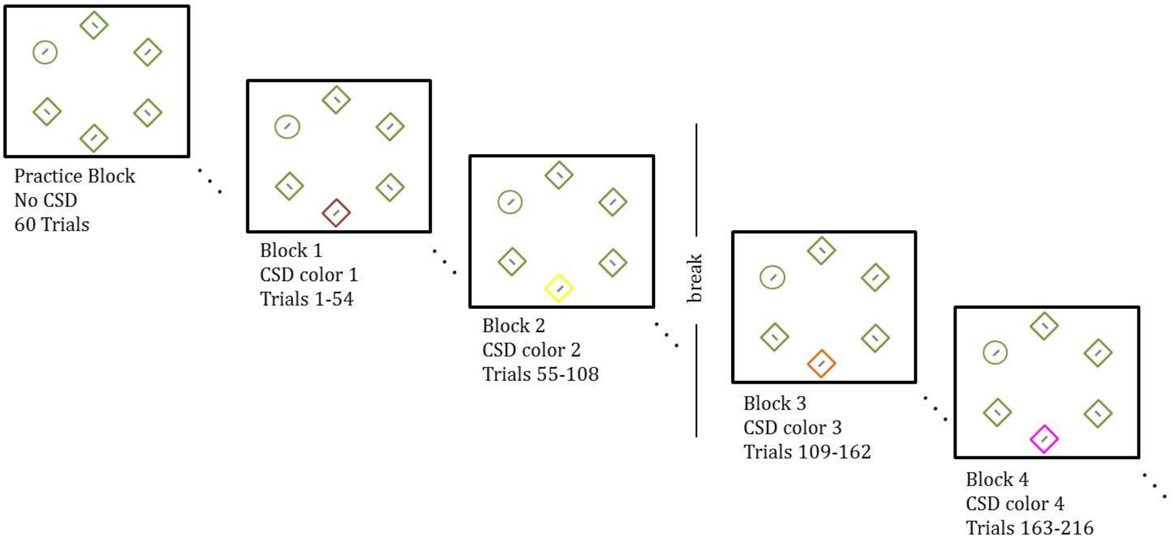

Figure 9. Schematic illustration of the trial procedure in the additional singleton task. For the sake of simplicity, only displays with set size six in the singleton display group are displayed. CSD position was of course randomly varied across trials (Study 2).

5.2.6 Overview of Outcome Variables

As eye movements were not recorded, response latencies and accuracy were the only dependent variables.

5.2.7 Data Preparation

Like in Study 1, the logarithm of the response latencies was computed and used as the dependent variable in all tests, whereas untransformed values are displayed in text, figures and tables. Trimming criteria were identical to the ones in Study 1. The datasets of two participants (both of them were in the feature display group) were excluded from analysis because their mean accuracy scores were at chance level (47% and 53%) and categorized as outlier values in a boxplot analysis. Less than 1% of the trials were discarded because they were outliers on response latencies. 4.4% of the responses were erroneous.

5.3 Results

5.3.1 The Effect of Set Size

The design of this study comprised four within-subjects variables (set size, trial type, block and sub-block), resulting in a total of 48 experimental conditions. Consequently, the number of trials within each condition was low, which did not allow analyzing the impact of all of the independent variables in one single ANOVA. For this reason, the effect of set size and the effects of block and sub-block were analyzed in separate ANOVAS. In this chapter, a mixed-design ANOVA on mean response time and mean accuracy with display group (feature display group, singleton display group), set size (six, nine or 12 items) and trial type (CSD absent, CSD present) as factors is reported.

5.3.1.1 Accuracy

The 2 (display group) x 3 (set size) x 2 (trial type) - ANOVA on accuracy revealed a significant interaction between trial type and set size, $F(2, 40) = 3.84$, $p = .03$, $\eta^2_p = .16$. In trials with set size nine, participants made more errors in CSD present trials, but in trials with set size 12, they made more errors in CSD absent trials (see Table 4). All of the other effects did not reach significance, *Fs* < 2.89, *ps* > .06.

Table 4. *Mean Accuracy in % (SD in Brackets), Dependent on Set Size and Trial Type (Study 2).*

	Accuracy	
	CSD absent	CSD present
Set size 6	96 (5)	95 (6)
Set size 9	97 (5)	95 (6)
Set size 12	95 (6)	96 (7)

5.3.1.2 Response Latencies

The 2 (display group) x 3 (set size) x 2 (trial type) - ANOVA on response latencies yielded a significant effect of trial type, $F(1, 20) = 45.03$, $p < .001$, $\eta^2_p = .69$ and a significant linear trend of set size, $F(1, 20) = 33.84$, $p < .001$, $\eta^2_p = .63$. Both effects were qualified by interactions with display group: Trial Type x Display Group, $F(1, 20) = 7.15$, $p = .015$, $\eta^2_p = .26$ and Set Size x Display Group, $F(1, 20) = 10.90$, $p = .004$, $\eta^2_p = .35$. Attentional capture was larger in the singleton display group ($M = 55$ ms) than in the feature display group ($M = 25$ ms), $t(20) = 2.67$, $p = .015$, 95% CI [4, 56]. Response latencies increased with set size in the feature display group [$F(1,9) = 26.56$, $p = .001$, $\eta^2_p = .75$], and in the singleton display group [$F(1,11) = 5.41$, $p = .04$, $\eta^2_p = .33$], with the search slope being steeper in the former. Search slopes were 12 ms/item in the feature search group and 2 ms/item in the singleton search group. The main effect of group and the interactions between trial type and set size and between trial type, set size and display group were non-significant, *F*s < 1 (see Table 5).

Table 5. *Mean Response Latencies in ms (SD in Brackets) for the two Display Groups, Dependent on Set Size (Study 2).*

	Feature display group		Singleton display group	
	CSD absent	CSD present	CSD absent	CSD present
Set size 6	712 (133)	736 (133)	717 (90)	776 (95)
Set size 9	759 (158)	784 (160)	719 (88)	788 (114)
Set size 12	779 (147)	806 (145)	741 (90)	778 (98)

5.3.2 The Effects of Sub-Block and Block in the Complete Sample

In this chapter, the effects of sub-block and block are analyzed with 2 (display group: feature display, singleton display) x 2 (trial type: CSD absent, CSD present) x 2 (sub-block: first, second) x 4 (block: 1 - 4) mixed-design ANOVAs. Thus, results were collapsed across different set sizes in order to reduce the number of conditions.

5.3.2.1 Accuracy

The 2 (display group) x 2 (trial type) x 2 (sub-block) x 4 (block) - ANOVA on accuracy did not show any significant effects, Fs < 4.24, ps > .05. Mean accuracy was 96% in both trial types.

5.3.2.2 Response Latencies

The 2 (display group) x 2 (trial type) x 2 (sub-block) x 4 (block) - ANOVA on response latencies revealed a main effect of trial type, $F(1, 20) = 41.33$, $p < .001$, $\eta^2_p = .67$, which was modulated by display group, $F(1, 20) = 5.16$, $p = .034$, $\eta^2_p = .21$. There were main effects of block, $F(3, 60) = 2.89$, p=.043, $\eta^2_p = .13$, and sub-block, $F(1, 20) = 9.25$, $p = .006$, $\eta^2_p = .32$. The main effect of trial type was also qualified by interactions with block, F (3, 60) = 6.40, $p = .001$, η^2_p = .24 and sub-block, $F(1, 20) = 13.18$, $p = .002$, $\eta^2_p = .40$. The other effects were not significant, Fs < 2.73, ps > .05. Although the three-way interactions between trial type, sub-block and display group and between trial type, block and display group were not significant [the latter was close to significance, $F(3, 60) = 2.72$, $p = .052$], separate ANOVAs for the two display groups were computed because observers in the feature display group applied serial search (the search slope was 12 ms/item, see Chapter 5.3.1.2) and because the focus of this study was whether attentional capture in the singleton display group is reduced by practice with the CSD color.

5.3.3 The Effects of Sub-Block and Block in the Feature Display Group

A 2 (trial type) x 2 (sub-block) x 4 (block) – ANOVA on response latencies in the feature display group resulted in a main effect of trial type, $F(1, 9) = 27.01$, $p = .001$, $\eta^2_p = .75$. This effect was qualified by the linear trend of block, $F(1, 9) = 10.53$, $p = .01$, $\eta^2_p = .54$, with lower interference in later blocks. In addition, the three-way interaction between trial type, sub-block and the linear trend of block was significant, $F(1, 9) = 6.51$, $p = .03$, $\eta^2_p = .42$. The remaining effects were not significant, *F*s < 2.80, *p*s > .12. As can be concluded from Figure 10, the CSD only delayed response latencies in the first sub-block of block 1.

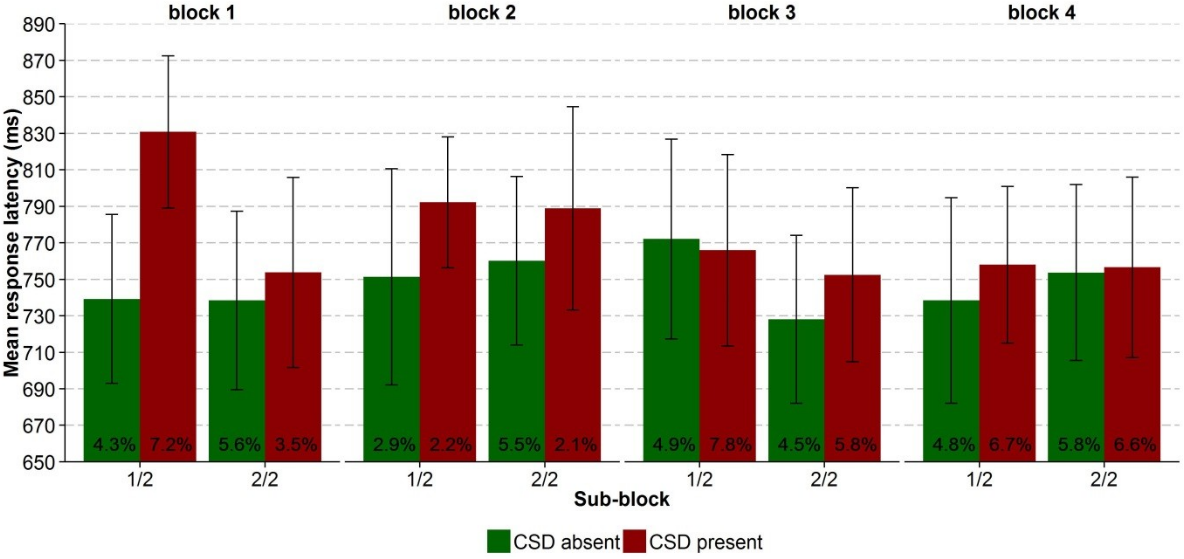

Figure 10. Mean response latencies and error rates (at the bottom of the bars) in the feature display group, dependent on sub-block, block and trial type. Error bars represent standard errors (Study 2).

5.3.4 The Effects of Sub-Block and Block in the Singleton Display Group

The 2 (trial type) x 2 (sub-block) x 4 (block) – ANOVA on response latencies in the singleton display group revealed a main effect of trial type, $F(1, 11) = 26.38$, $p < .001$, $\eta^2_p = .71$ and of sub-block, $F(1, 11) = 7.93$, $p = .017$, $\eta^2_p = .42$. As expected, the effect of trial type was modulated by sub-block, $F(1, 11) = 40.31$, $p < .001$, $\eta^2_p = .79$. Furthermore, the interaction between trial type and block was also significant, $F(3, 33) = 6.40$, $p = .002$, $\eta^2_p = .37$. The remaining effects were not significant, *F*s < 2.64, *p*s > .06. To disentangle the interaction between trial type and sub-block, response latencies in CSD present and CSD absent trials in each sub-block were compared with dependent *t*-tests to test whether attentional capture occurred in each sub-block. These revealed that the CSD only delayed responses in the first sub-block of block 1 and block 2. In the third block, the CSD did not slow reaction time at all. Unexpectedly, attentional capture occurred again in both sub-blocks of block 4 (see Table 6 and Figure 11).

In order to further examine the interaction between trial type and block, the mean amount of attentional capture for each block was compared between blocks. *T*-tests revealed that attentional capture increased significantly between the first and second block, but decreased significantly from block 2 to block 3 and then did not change again between block 3 and block 4. Attentional capture in block 4 did not significantly differ from attentional capture in in block 1 and block 2 (see Table 7 and Figure 11).

Table 6. *Attentional Capture in the Singleton Display Group, Dependent on Block and Sub-Block (Study 2).*

		Attentional Capture			95% CI	
		M (SD)	*t(11)*	*p*	*LL*	*UL*
Block 1	Part 1	85 (86)	3.72	.003	30	139
	Part 2	27 (72)	1.07	.306	-18	73
Block 2	Part 1	128 (90)	4.88	<.001	71	185
	Part 2	50 (96)	2.05	.066	-11	110
Block 3	Part 1	38 (69)	1.92	.081	-7	82
	Part 2	10 (46)	0.52	.614	-19	39
Block 4	Part 1	48 (49)	3.54	.005	17	79
	Part 2	44 (47)	3.84	.003	14	74

Note. Attentional Capture = difference between mean response latency in CSD present trials and mean response latencies in CSD absent trials. LL = lower level of CI; UL = upper level of CI. *t*- and *p*-values are derived from *t*-tests with the logarithmized response latencies; *M*, *SD* and 95% CI indicate the untransformed response latencies in ms.

Table 7. *Attentional Capture Difference Scores in the Singleton Display Group (Study 2).*

	Attentional Capture Difference Score			95% CI	
	M (SD)	*t(11)*	*p*	*LL*	*UL*
Block 1 – Block 2	-33 (59)	-2.75	.019	-70	4
Block 2 – Block 3	65 (53)	4.49	.001	32	99
Block 3 – Block 4	-22 (46)	-1.99	.072	-51	7
Block 1 – Block 4	9 (59)	-0.29	.778	-28	47
Block 2 – Block 4	43 (59)	2.10	.059	5	80

Note. Attentional Capture Difference Score = difference in attentional capture between two blocks. LL = lower level of CI; UL = upper level of CI. *t*- and *p*- values are derived from *t*-tests with the logarithmized response latencies; *M*, *SD* and 95% CI indicate the untransformed response latencies in ms.

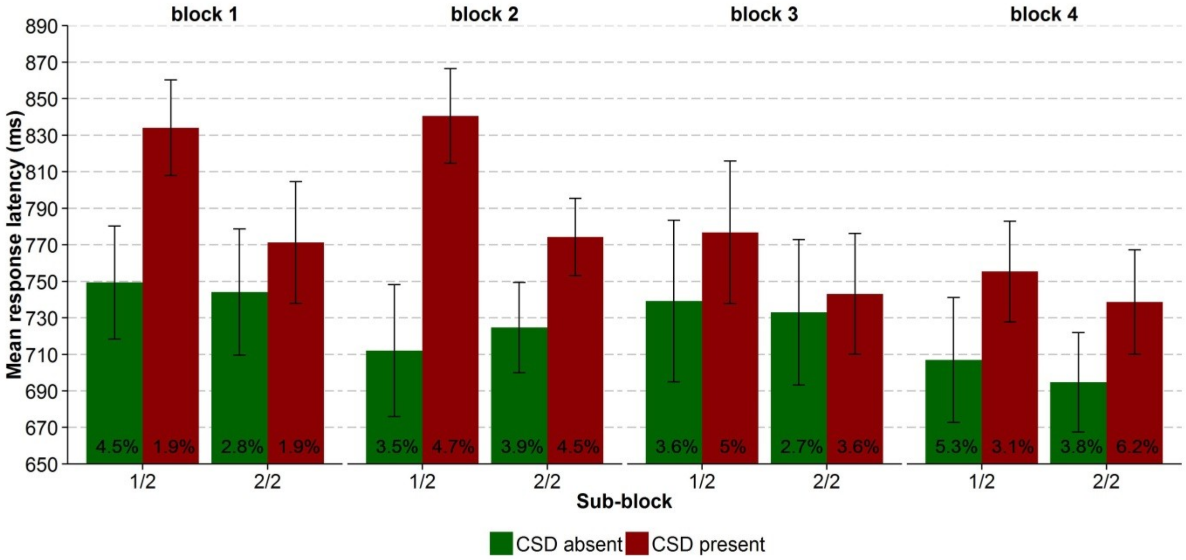

Figure 11. Mean response latencies and error rates (at the bottom of the bars) in the singleton display group, dependent on sub-block, block and trial type. Error bars represent standard errors (Study 2).

5.3.5 The Effect of Visual Working Memory Capacity on Attentional Capture

K was calculated as the indicator of visual working memory capacity with the following formula: $K = S \times (H - F)$ (Cowan, 2001; Fukuda & Vogel, 2009, 2011). S is the number of items in the display, H is the hit rate (the number of trials in which a change was correctly identified divided by the number of trials with a change) and F is the false alarm rate (the number of trials in which a change was erroneously detected divided by the number of trials without a change). K ranged from 1.9 to 4.8 (M = 3.3 in the singleton display group, M = 3.6 in the feature display group). Feature and display group did not differ significantly in K, $t(20) = -1.09$, $p = .29$ and K did not correlate significantly with the mean amount of attentional capture (mean response latency in CSD present trials – mean response latency in CSD absent trials) during the first half of the first block, $r(22) = .17$, $p = .44$. Consequently, all further analyses with K were skipped because it did not explain any variation between participants in their initial amount of attentional capture.

5.4 Discussion

The aim of Study 1 was to test the hypothesis that attentional capture in the additional singleton task with homogeneous distractors is attenuated very quickly by distractor practice and therefore only occurs in the first half of a block. Furthermore, set size was manipulated in this study to verify that visual search was parallel in Study 1 and that serial search is not an explanation for the absence of attentional capture in Study 1. The analysis in Chapter 5.3.1.2 demonstrated that the effect of set size on response latencies was significant in both display groups. The search slope in the singleton display group, which searched among exactly the same display as the participants in Study 1, however, was very small (2 ms/item) and well below the cut-off of 10 ms/item (see Chapter 2.1.1). This result verifies that search in the additional singleton task with homogeneous distractors, which was employed in Study 1, was parallel. Any interference effect measured in the singleton display group must thus be due to attentional capture. The search slope in the feature display group, however, was 12 ms/item and indicates serial search. Therefore, interference in response latencies in this group cannot be interpreted as a sign of attentional capture because search was not parallel.

In both display groups, the main effect of trial type, collapsed across all of the other within-subject variables, was large (η^2_p = .75 in the feature display group, η^2_p = .71 in the singleton display group). Vatterott and Vecera (2012) were the first to include block as a factor in the analysis of CSD effects. This result, in combination with the findings of Vatterott and Vecera (2012), suggests that the reported main effects of attentional capture in the literature deliver only an incomplete picture of the process of the distraction effect of the CSD over the course of the experiment. Mean interference was 25 ms in the feature search group and 55 ms in the singleton search group, yet participants in the former group were only distracted in the first twenty-seven trials of the experiment and the latter showed marked (and partly unexpected) changes in the course of attentional capture, which poses the question what attentional capture as a main effect really means. The next two chapters discuss the results for the feature display group (see Chapter 5.4.1) and the singleton display group (see Chapter 5.4.2). Chapter 5.4.3 addresses some limitations of this study and the last chapter (5.4.4) provides a conclusion.

5.4.1 Distractor Interference in the Feature Display Group

Participants in the feature display group searched for the target among heterogeneous distractors. Analysis showed that observers in this condition were distracted by the CSD in the first part of block 1, but not later on. This pattern of results suggests that participants learned very quickly to suppress the CSD and were not prone to distraction, even when its color changed. Nevertheless, the main effect of trial type was significant, even when the data were aggregated over the four blocks of the experiment – as it is typically done in studies working with this paradigm. Thus, Study 2 showed that distractor interference in feature search mode is partly determined by practice with the CSD, which is consistent with the results of Vatterott and Vecera (2012). Although the participants in the feature display group practiced feature search in the practice block over 60 trials, they were very distracted by the appearance of the CSD in the first test block and showed an overall distractor interference of 25 ms. This effect might at first hand suggest that feature search mode could not have been at play here – if one assumes that feature search mode prevents any form of attentional capture. Studies that manipulated search modes, however, used 480 training trials (Leber & Egeth, 2006; Zehetleitner et al., 2012) to induce the respective search mode. Zehetleitner et al. (2012) reported a significant attentional capture effect of 10 ms in the training phase in the feature search group. Based on the present results, the interference found by Zehetleitner et al. could be based on large interference in the beginning of the training phase, which is then attenuated (the authors did not report block-wise analysis). Therefore, a distraction effect of the CSD in the first sub-block of block 1 is not that surprising and is still compatible with feature search mode because participants obviously need more than 60 trials to acquire an effective search mode strategy. In addition, the absence of CSD interference from the second block on supports the explanation that participants adopted feature search mode.

Yet, the general decrease of distractor interference over the course of the blocks is clearly at odds with the result of Vatterott and Vecera (2012). One significant difference in design between Study 2 and the study by Vatterott and Vecera is that set size was varied in the former. The search slope of 12 ms/item in the feature display group indicates serial search and is comparable to that reported by other studies employing heterogeneous distractors and varying set size (Bacon & Egeth, 1994; Theeuwes, 2004). Vatterott and Vecera presumed that search was parallel in their study, but they did not vary set size or record eye movements to prove it. Support for their view comes from studies that observed flat search slopes with similar stimuli (Leber & Egeth, 2006; Zehetleitner et al., 2012). Thus, the evidence concerning search slopes with heterogeneous distractors is mixed: sometimes they lead to parallel and sometimes they lead to serial search. Therefore, one obvious explanation for the discrepancy in results between Study 2 and the study by Vatterott and Vecera (2012) is that search was serial in the former and parallel in the latter. As a consequence, the reduction of distractor interference was larger in Study 2 because serial search enabled participants to control the saliency of the distractor, for example, by adopting a focused attentional window. (Theeuwes, 2004; see Chapter 2.1.3) .

Another explanation for the conflicting results is based on the stimulus eccentricity. The distance between the center of the screen and the center of the stimuli was 11.4° in Study 2 (in order to make eye movements necessary for successful search) and only 4.2° in the study by Vatterott and Vecera (2012). Whereas set size ranged from six to 12 stimuli in Study 2, only six stimuli were employed in the latter study. As the size of the stimuli was the same, the display was denser in the study by Vatterott and Vecera (2012) than in set size 6 of Study 2 (the higher number of stimuli with set size nine and 12 compensates for eccentricity). Density affects visual saliency: The higher the density of the display, the higher the salience of a singleton distractor and of a target in the display (Yantis & Egeth, 1999). Thus, stimulus eccentricity was confounded with visual density and visual saliency in one condition (set size 6) of Study 2. Were participants able to reduce distractor interference to a greater degree in Study 2 because the CSD was less salient at set size six? If that was the case, interference should be reduced with six stimuli compared to conditions with nine or 12 stimuli. The data, however, did not confirm this assumption. Set size, trial type and display group clearly did not interact in the analysis of response latencies, $F < 1$. An additional ANOVA on response latencies in the feature display group revealed a non-significant interaction between trial type and set size, $F < 1$. In addition, it is important to note that Study 2 revealed a large main effect of trial type in the feature display group. In the study by Vatterott and Vecera, however, the main effect of trial type was only marginally significant, $F(1, 15) = 3.62$, $p < .08$. The amount of attentional capture/distractor interference in the first sub-block was similar in both studies (105 ms in the study by Vatterott and Vecera and 92 ms in Study 2). This pattern of results suggests that the salience of the CSD was comparable in both studies.

To conclude, the discrepancy in results is best explained with reference to the differences between parallel and serial search. The previous argumentation suggests that a target can be detected among heterogeneous distractors in parallel search if the stimulus eccentricity is low and if only eye movements of short amplitude are required (cf. Vatterott & Vecera, 2012), whereas a target among heterogeneous distractors, which can only be detected via eye movements to the periphery of the display, induces at least partly serial search (as in Study 2), which facilitates distractor suppression over the course of the trial.

5.4.2 Attentional Capture in the Singleton Display Group

Observers in the singleton display group had to search for the target among a set of homogeneous distractors. As search was parallel in this group, attentional capture is the obvious interpretation of any interference by the CSD. As expected, results showed that practice with the specific distractor feature is a sufficient precondition for effective distractor suppression, even when participants are confronted with a display that only includes homogeneous distractors. Attentional capture dropped to a non-significant level in each of the second sub-blocks with the exception of block 4. One might speculate that, if the distractor had not changed its color after each block, the main effect of trial type would not have been significant. This interpretation is at odds with the stimulus-driven capture account (Theeuwes, 2010), but in accordance with the dimension weighting account (Müller et al., 2009; Zehetleitner et al., 2012). In the study by Müller et al. (2009, Exp. 2) and in Study 1, the CSD did not switch color and attentional capture was close to zero with 50% CSD probability. Therefore, Study 2 provides an important modification to the findings of Vatterott and Vecera (2012), which suggested that the fast distractor practice effects are dependent on feature search mode.

Unlike in the study by Vatterott and Vecera (2012), which reported stable distractor interference in each of the four first sub-blocks, the amount of attentional capture varied considerably between the blocks in the present study. It increased from the first to the second block, then decreased again in block 3 and remained unchanged in block 4. Mean interference was larger in block 4 than in block 3, but this difference was not reliable (p = .072). On the other hand, mean interference was decreased in block 4 compared to block 2, but this difference also failed to reach significance, p = .059. Nevertheless, this result might be indicative of a general practice effect, with lower attentional capture in later blocks. The dimension weighting account (Zehetleitner et al., 2012) assumes, in contrast to the search mode account (Bacon & Egeth, 1994), that distractor suppression is not only dependent on specific distractor practice, but also on general distractor practice. Learning to suppress one specific distractor feature (e.g., blue) makes it easier for participants to learn to suppress another specific distractor feature (e.g., yellow) (Zehetleitner et al., 2012). Distractor suppression is transferred from one feature to another, and specific practice with a feature becomes less important the more practice one has acquired with other features previously. The authors explained this learning effect with hierarchical feature weighting. The feature weights are interdependent in their respective dimension. For example, if the feature blue in the dimension color is reduced (because one knows that the distractor is blue), this also leads to a reduction of the feature weights of the other colors.

This account can explain the - non-significant - decrease in capture between block 2 and block 4, but it cannot explain the increase in capture from block 1 to block 2. Block 1 and block 2 differ in the expectancy that they invoke in the observers: At the beginning of block 1, observers expected to encounter a CSD, but they did not know its color. The beginning of block 2, however, was not explicitly communicated to participants. They knew that the CSD would change its color, but did not know when exactly. Therefore, observers might have been surprised by the sudden appearance of a differently colored CSD in block 2. Horstmann (2005) has shown that a CSD captures attention only if it deviates from expectancy. Expectancy was manipulated in his experiments by varying the CSD color in precritical trials: These could either contain distractors with one color (e.g., only green) or two colors (e.g., green and white). In the former condition, participants assumed to expect only one distractor color over the course of the experiment, whereas in the latter condition, participants formed the expectancy to encounter distractors of various colors. Horstmann compared response latencies in the precritical trials (48 trials), in the critical trial (trial 49) and in the post-critical trials. If only one distractor color was presented in the pre-critical trials, attention was captured in the critical trial, whereas interference was not observed in the critical trial if the pre-critical trails contained two distractor colors. Horstmann argued that the new color violated the expectancy of the observers.

In contrast to Study 2, Horstmann did not inform the participants of a possible change in distractor color during the experiment. Still, one could argue that the appearance of the CSD in block 2 was different to the appearance of the CSD in block 1: The beginning of block 1 was clearly marked, whereas the transition from block 1 to block 2 was not introduced by any message. For the same reason, the appearance of the CSD in block 3 might have been less unexpected than the appearance of the CSD in block 2 and block 4: Block 3 was clearly separated from block 2 by the short break in the middle of the experiment. In block 4, however, the amount of attentional capture was marginally significantly lower than in block 2, although the CSD appeared as unexpectedly as in block 2. Thus, the increase in response latencies from block 1 to block 2 and the tendency of a decrease in attentional capture from block 2 to block 4 suggest that, at the beginning of the experiment, the expectancies of the observers affected the amount of capture, but that later on, they learned to suppress even unexpected distractors due to general distractor practice.

Another pattern in the data differs from the results reported by Vatterott and Vecera (2012). In Study 2, the CSD did not capture attention in both sub-blocks of block 3, but delayed response times in both sub-blocks of block 4. Participants' expectancy to encounter a new CSD color after the break and general distractor practice might have facilitated distractor suppression at the start of block 3. Distractor suppression might not have work as well in block 4 as in block 1 and block 2 due to a decrease in response latencies: Attentional capture is higher with faster response latencies (Siebold et al., 2011; van Zoest et al., 2004). Response latencies in block 4, however, were not significantly reduced compared to response latencies in block 1, $t(11) = 1.95$, $p = .077$, or in block 3, $t(11) = 1.39$, $p = .19$. A speculative idea, which cannot be corroborated by the data, is that participants' motivation to exert attentional control during search was depleted by the end.

To sum up, the results in the singleton display group confirmed Hypothesis 2.1 because attentional capture was modulated by fast CSD practice and disappeared after 27 trials. Nevertheless, some results were surprising: This discussion could not provide a definite answer to the question why attentional capture was absent in block 3 and then reoccurred in both sub-blocks of block 4. An interesting finding is the marginally significant decrease of attentional capture over the course of the experiment. Based on these results, the most likely explanation for the absence of attentional capture in the neutral condition in Study 1 is that participants learned very quickly to suppress the CSD in the practice block and therefore were no longer distracted by it in the test trials.

5.4.3 Limitations

Some arguments presented in the previous chapters have already hinted at limitations of the present study. Firstly, one cannot be absolutely certain that individuals in the different display groups actually applied different search modes. The most recent studies examining search modes consisted of two phases: In a training phase of 480 trials, participants were induced to use either feature or singleton search mode. Participants in the feature search mode searched for a constant target among heterogeneous distractors, whereas observers in singleton search mode searched for a variable target among homogeneous distractors. Afterwards, participants of both groups completed 480 test trials with the same display (constant target, homogeneous distractors) (Leber & Egeth, 2006; Zehetleitner et al., 2012). This elegant design was not chosen for this Study due to time restrictions for data collection. Instead, the same method as in the study by Vatterott and Vecera (2012) was used. Therefore, the data do not allow a definite conclusion as to what was the cause for the different response pattern in the feature display group, compared to the singleton display group: was it the search mode induced by the feature search display or the serial search process?

A second limitation concerns the analysis: The four within-subjects variables that were included in the design (trial type, set size, sub-block and block) could not be analyzed in one single ANOVA because the number of trials in each condition was too low. If one had wanted to include all of the variables into one analysis, the total number of trials in each block would have needed to be much higher than in the study by Vatterott and Vecera (2012), which used 48 trials per block, and the comparison with this study would have been limited. Furthermore, there is no reason to assume that set size effects vary over the course of the experiment.

A third limitation pertains to the sample size. The sample size in Study 2 (ten, respectively twelve participants in each group) was similar to that of almost any other study on attentional capture (Bacon & Egeth, 1994; Geyer et al., 2008; Leber & Egeth, 2006; Müller et al., 2009; Theeuwes, 1991, 1992, 2004; Theeuwes et al., 2003; Wienrich & Janczyk, 2011; Zehetleitner et al., 2012). Therefore, Study 2 had the same power as those studies. The power was, however, probably not sufficient to detect individual differences in attentional capture, depending on visual working memory.

5.4.4 Conclusion

The results of Study 1 were rather inconclusive regarding the effect of incentives on goal-directed control over attentional capture in the additional singleton task. Consequently, it was the aim of Study 2 to study the role of distractor practice in the additional singleton task without manipulating reward, in order to find out if the task is adequate to study the moderating influence of incentives. As this doctoral thesis aims to study covert and overt attention, a stimulus eccentricity has to be used that forces participants to make eye movements to detect the target. The results in the singleton display group are especially interesting because its display closely resembles the one used in Study 1. With this display, search is clearly parallel (although response latencies did increase with set size, the slope was close to zero) and any influence of a feature search mode is highly unlikely. Therefore, this display is adequate to examine the influence of distractor practice and the incentive to use distractor suppression as outlined by the dimension weighting account (Müller et al., 2009; Zehetleitner et al., 2012).

The central finding of this study - the fast decrease of attentional capture in singleton search mode - has important implications for the design of the search task and the payoff scheme in the next study. In the singleton display group, attentional capture decreased to a non-significant level in the second sub-blocks of block 1 and 2, was totally absent in block 3 and reappeared in both sub-blocks of block 4. If one maps the data pattern expressed in block 1 and block 2 to an experiment in which the CSD has only one color, it is speculative, but not unlikely, that attentional capture is only observable in the first 30 trials of the first block. This leaves only thirty trials in which to investigate the influence of incentives in a within-subjects design. Therefore, changing the distractor color with each block seems to be a promising method to induce large interference effects anew several times in the course of an experiment. This would result in roughly 120 (4 x 30) trials with attentional capture, in which the effect of incentives can be studied. Furthermore, the present results indicate that attentional capture also decreases over the course of the blocks, even if the color of the CSD switches after each block. To ensure a large amount of attentional capture throughout several experimental blocks, a different variant of the additional singleton task was employed in Study 3. In his first study on attentional capture in the additional singleton task, Theeuwes (1991) used variable targets: the target and the distractors switched forms from trial to trial (see Chapter 2.1.3). With this task, he observed an average interference effect of about 120 to 150 ms, whereas the average amount of attentional capture with a constant target is approximately 20 ms (cf. Theeuwes, 1992). Finally, as Study 2 has shown that distractor interference decreases linearly from the beginning to the end of a block, incentive was not manipulated trial-wise, but block-wise.

6 STUDY 3: THE EFFECT OF REWARD ON MENTAL EFFORT AND DISTRACTOR INTERFERENCE

6.1 Hypotheses

The aim of Study 3 was to use the results of Study 2 in order to investigate the hypotheses of Study 1 with a more adequate search task and payoff scheme. Therefore, the hypotheses of Study 3 were almost the same as the ones of Study 1.

- **Hypothesis 3.1:** Performance-contingent monetary incentives increase the mental effort invested during the additional singleton task, compared to a baseline condition with no incentives.
- **Hypothesis 3.2:** Performance-contingent monetary incentives increase the search performance, compared to a baseline condition with no incentives.
- **Hypothesis 3.3:** Performance-contingent monetary incentives decrease covert and overt capture by the CSD, compared to a baseline condition with no incentives.

6.2 Method

6.2.1 Participants

Twenty-seven students (16 female) of the University of Heidelberg took part in the experiment. Their age ranged from 19 to 37 years (*M* = 24). The participants received the amount of money that they won in the experiment as compensation.

6.2.2 Design

Trial type, *incentive* and *sub-block* were manipulated within subjects in this experiment: First, half of the trial trials contained a CSD (*CSD present* trials) and the other half did not (*CSD absent* trials). Second, there were two different incentive conditions which varied block-wise: in *reward* blocks, participants won 10 ct in 80% of the trials if they made a correct and fast response. In each reward block, 80% of the trials (i.e., 48 trials) were randomly determined to be reward trials, in which participants could potentially receive 10 ct. Like in Study 1, the correct response had to be within an individual and adaptive time limit - the median of the last six responses plus an additional 50 ms. If participants made an incorrect response, they lost 10 ct.

Thus, in contrast to Study 1, errors were explicitly punished in order to motivate participants to avoid errors. In neutral blocks, participants could not win any money. In contrast to Study 1, there was no loss condition. Therefore, the number of reward trials was limited in order to ensure that participants would receive, on average, an amount of reward that would be similar to the standard monetary fee for participation in an experiment that took approximately 30 minutes.

In analysis, each block was split in two halves of 30 trials because attentional capture was expected to be only present in the first sub-block of each block. Like in Study 2, the CSD color changed in each block. The sequence of CSD colors and of reward and neutral blocks was counterbalanced across participants. Participants either started with a neutral block (*neutral start group*) or with a reward block (*reward start group*) after which neutral and reward blocks alternated. Their initial account balance was 0.00 €. Like in Study 1, they were told that they would only receive the amount of money that they won based on their performance. After the experiment, however, each participant received at least 3.00 €.

6.2.3 Stimuli

In each trial, six stimuli were presented on white (RGB: 255, 255, 255) background[4] on an imaginary circle with a radius of 11.4°. The target was either a circle (diameter: 2.3°) among diamonds (a square with a side length of 1.9° rotated by 45°) or vice versa. Size and form of these objects was the same as in Study 1 and in Study 2. All objects had a line width of 0.1°. All of the stimuli were green (RGB: 114, 201, 101) except for the CSD, which was red (RGB: 255, 144, 144), blue (RGB: 173, 173, 255), yellow (RGB: 180, 180, 0) or pink (RGB: 255, 132, 155), depending on block and sequence group. Target lines were identical to the ones used in Study 1. The RGB colors of the stimuli were taken from Wentura, Müller, and Rothermund (2013), who matched objective luminance of the colors by measuring luminance via a luminance meter.

6.2.4 Apparatus

The apparatus was the same as in Study 1, except that manual responses were recorded with highly accurate response pads from a different manufacturer (RB-740 by Cedrus Corporation).

---

[4] The mirror of the eye tracker was reflected on the black background of the computer screen in Study 1. Therefore, the background color was changed to white in Study 3 in order to prevent light reflections.

6.2.5 Procedure

After participants had signed an informed consent and had read the instructions, the eye tracker was calibrated. 13 small squares appeared at different positions on the screen one after another and had to be fixated by the observer. In case of a successful calibration, the experimental task started. If recorded gaze position fell within a radius of 50 pixels around the center of the fixation cross, and the fixation cross was already presented on screen for at least 500 ms (in order to control for anticipatory fixations), the fixation was regarded as valid. In case of a valid fixation, the fixation cross remained on screen for another 500 ms before the search display was presented. If a valid fixation did not fall on the fixation cross within 2000 ms after its initial presentation, a recalibration was initiated. This fixation check was added to the experimental procedure to ensure that a trial only started if the image analysis software of the eye tracker was able to track the gaze position correctly. Then, the search display was presented until the observer responded. The composition of the display was different to Study 1 and Study 2 because the more difficult version of the additional singleton task (Theeuwes, 1991) was used: The target was a circle among diamonds in 50% of trials and a diamond among circles in the rest of the trials. Thus, search was more difficult in comparison to Study 1 and Study 2 because participants did not know the exact target features, but had to search for the discrepant green item. Like in the previous studies, participants knew that any item that was not green could be ignored. Participants' task was to press the right response key if the line inside the target (the discrepant form) was tilted to the right (/) and the left response key if the line inside the target was tilted to the left (\). They were informed that in some trials one of the distractors would be of a different color than green and that they should ignore it. Instructions equally emphasized speed and accuracy.

Immediately after their response, participants received feedback on their performance. If they gave the wrong response, the word "Fehler" (error) was presented. If their response time exceeded their individual time limit for a fast response, the words "zu langsam" (too slow) appeared. In addition, the current reward (+10 ct, in green color) or loss (-10 ct, in red color) and the current account balance were presented in the middle of the screen. Feedback remained on screen for 1000 ms. The next trial started after an intertrial interval of 500 ms. Participants completed one practice block of 20 trials and continued with four blocks of 60 trials each. The eye tracker was calibrated at the beginning of each block, giving participants the opportunity to have a short break. One session took approximately 30 minutes.

6.2.6 Eye Movement Event Detection

The same criteria as in Study 1 were applied.

6.2.7 Overview of Outcome Variables

The same outcome variables as in Study 1 were analyzed. In addition to phasic pupil dilation, tonic pupil dilation was analyzed because incentives were manipulated block-wise.

6.2.8 Data Preparation

The same procedures as in Study 1 and Study 2 were applied. Errors made up 11.6% of all trials. 1.1% of the trials were categorized as outliers on response latency. 7.4% of the trials had to be excluded due to an anticipatory saccade (0 ms > saccadic latency < 80 ms), 1.8% due to a saccadic latency of 0 ms and 0.6% because the z-transformed saccadic latency exceeded 3.29. Overall, 22.5% of the trials had to be discarded due to these criteria. As a result, the datasets of two participants had to be excluded from analysis because they lacked data in at least one outcome variable in one cell of the conditions (30%, respectively 61% of their trials had to be excluded). The data of one participant was not recorded due to program failure. Thus, eleven participants remained in the neutral start and 13 participants in the reward start group.

6.3 Results

This chapter presents two types of analysis. The first analysis is based on the hypotheses and examines whether attentional capture is decreased in reward blocks compared to neutral blocks. Testing this hypothesis means that dependent variables are collapsed over the two reward and neutral blocks (see Chapter 6.3.1). The second chapter presents an exploratory analysis. This addresses the question whether the first analysis might blur practice effects over the course of the four blocks. Thus, the second analysis skips the variable incentive and replaces it with the variable block (see Chapter 6.3.2).

6.3.1 Analysis 1: Testing the Hypotheses

The first analysis was conducted with incentive (neutral block, reward block), trial type (CSD absent, CSD present), block (1 - 2) and sub-block (1 - 2) as within-subjects variables and sequence group (reward start group, neutral start group) as a between-subjects factor. This factor was added to the analysis in order to test sequence effects of reward. For the sake of brevity, the main focus rests on effects of trial type and incentive, so that not every significant effect that might be revealed by an ANOVA is described in detail. It is important to note that the variable block only contains two conditions (block 1 and block 2) because two of the four blocks were reward blocks and two were neutral blocks. Thus, “block 1” refers to the first reward or neutral block (the first or the second block in the sequence of the experiment, depending on the sequence group) and block 2 refers to the second reward or neutral block (the third or the fourth block in the sequence of the experiment, depending on the sequence group).

6.3.1.1 Amount of Rewards Won

Based on their actual performance, participants won 4.54 € (range [1.30, 6.30], *SD* = 1.24 €). The two sequence groups did not differ in the amount of money won, $t < 1$.

6.3.1.2 Accuracy

The 2 (incentive) x 2 (trial type) x 2 (block) x 2 (sub-block) x 2 (sequence group) - ANOVA on accuracy revealed a main effect of incentive, $F(1, 22) = 6.00$, $p = .023$, $\eta^2_p = .21$ and of trial type, $F(1, 22) = 6.42$, $p = .019$, $\eta^2_p = .23$, with fewer errors in reward blocks and in CSD absent trials. The main effect of trial type was qualified by an interaction with sequence group, $F(1, 22) = 12.82$, $p = .002$, $\eta^2_p = .37$. The three-way interaction between block, trial type and sequence group reached significance, $F(1, 22) = 5.98$, $p = .023$, $\eta^2_p = .21$. While participants in the neutral start group made more errors in CSD present trials in both blocks, [$F(1, 12) = 23.56$, $p <.001$, $\eta^2_p = .66$ in block 1 and $F(1, 12) = 6.31$, $p = .027$, $\eta^2_p = .34$ in block 2] participants in the reward start group tended to make more errors in CSD absent trials in block 1, $F(1,10) = 2.34$, $p = .16$, but did not show any interference in block 2, $F < 1$ (see Figure 12). The remaining effects were non-significant, *F*s < 2.62, *p*s > .11.

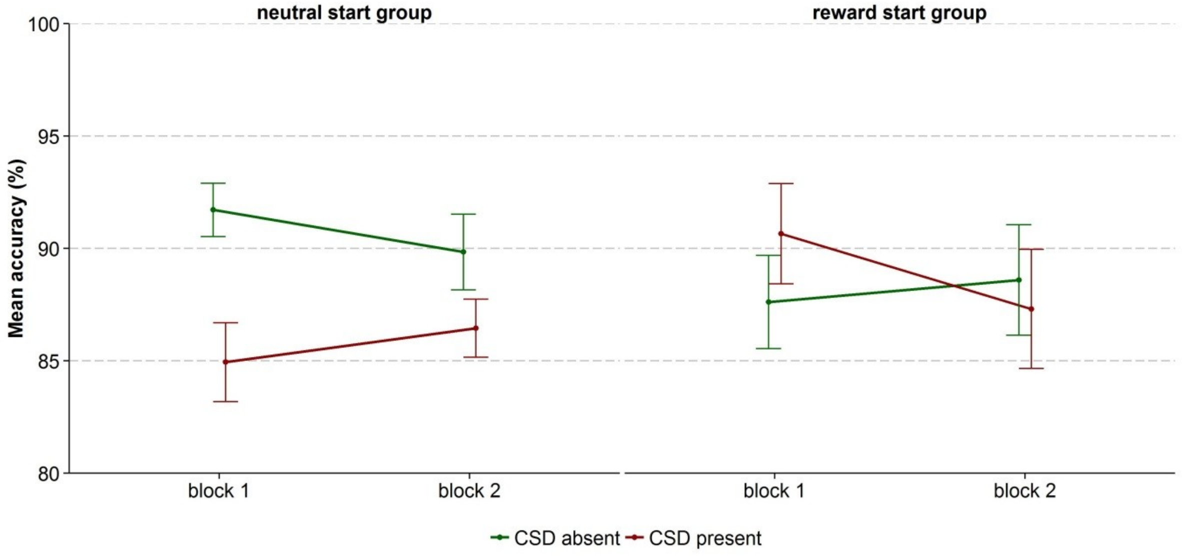

Figure 12. Mean accuracy, separately for the two sequence groups, dependent on block and trial type. Error bars represent standard errors. Points are set off horizontally to enhance the visibility of the error bars (Study 3).

6.3.1.3 Response Latency

The 2 (incentive) x 2 (trial type) x 2 (block) x 2 (sub-block) x 2 (sequence group) – ANOVA on response latency resulted in significant main effects of trial type, $F(1, 22) = 40.64$, $p < .001$, $\eta^2_p = .65$, block, $F(1, 22) = 53.13$, $p < .001$, $\eta^2_p = .71$, and sub-block, F(1, 22) = 16.55, $p = .001$, $\eta^2_p = .43$ with faster response latencies in CSD absent trials, second blocks and sub-blocks. The interaction between incentive and sequence group was significant, $F(1, 22) = 7.81$, $p = .011$, $\eta^2_p = .26$ and was further modulated by block, $F(1, 22) = 10.88$, $p = .003$, $\eta^2_p = .33$. In addition, the three-way interaction between incentive, trial type and block was significant, $F(1, 22) = 4.39$, $p = .048$, $\eta^2_p = .17$. Separate 2 (trial type) x 2 (block) x 2 (sub-block) x 2 (sequence group) – ANOVAs for the two incentive conditions revealed a significant Block x Trial Type interaction in reward blocks, $F(1, 22) = 7.26$, $p = .013$, $\eta^2_p = .25$, whereas this interaction was not significant in neutral blocks, $F < 1$. Thus, the amount of attentional capture decreased in reward blocks, but remained stable in neutral blocks. However, the amount of distractor interference in the second block did not differ between neutral and reward blocks, $F < 1$ (see Figure 13). The remaining effects did not reach significance, Fs < 3.65, ps > .06.

Thus, in contrast to Study 2, trial type and sub-block did not interact. The main effect of trial type was present both in response latencies and accuracy, showing that participants were faster to find the target and made fewer errors in CSD absent trials. Apart from this effect, none of the aforementioned effects was present in both manual response variables. Whereas a main effect of reward was not observed in response latencies, participants made fewer errors in reward blocks.

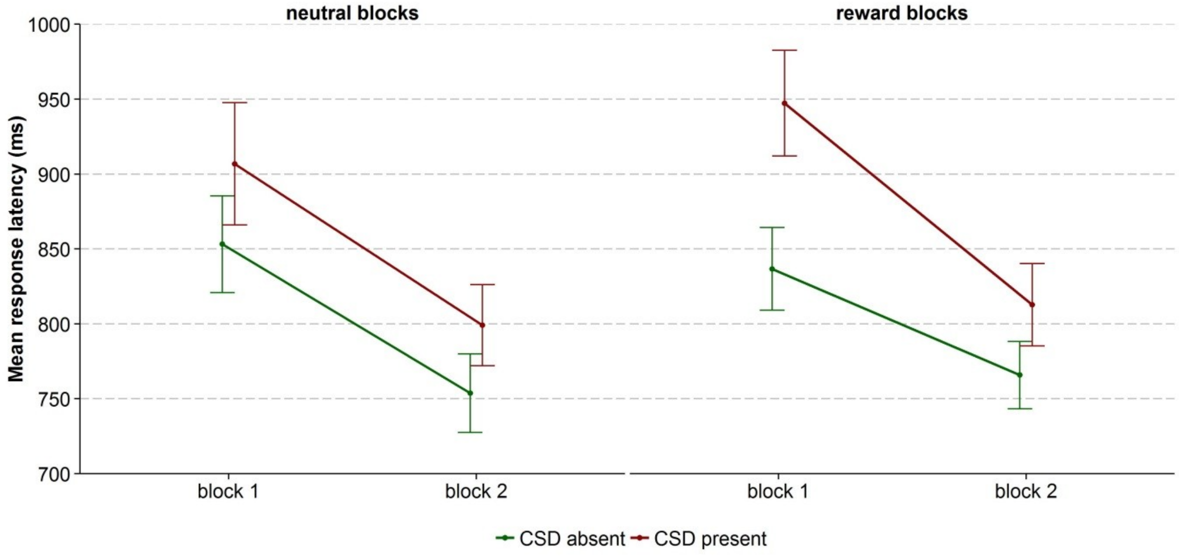

Figure 13. Mean response latencies, dependent on incentive and trial type. Error bars represent standard errors. Points are set off horizontally to enhance the visibility of the error bars (Study 3).

6.3.1.4 Tonic Pupil Dilation

Tonic pupil dilation refers to the mean pupil diameter during the presentation of the fixation cross at the beginning of the trial. This variable was added, in comparison to Study 1, as an indicator of mental effort because reward was manipulated block-wise. Therefore, participants' sustained mental effort, next to their transient mental effort (see Chapter 6.3.1.5), might have been affected by the rewards. The 2 (incentive) x 2 (trial type) x 2 (block) x 2 (sub-block) x 2 (sequence group) - ANOVA on tonic pupil dilation revealed main effects of block, $F(1, 22) = 37.64$, $p < .001$, $\eta^2_p = .63$, sub-block, $F(1, 22) = 59.40$, $p < .001$, $\eta^2_p = .73$, and incentive, $F(1, 22) = 34.39$, $p < .001$, $\eta^2_p = .61$, with larger pupil diameter in the first block and sub-block and in the reward blocks. The main effect of block was qualified by an interaction with sub-block, $F(1, 22) = 5.78$, $p = .025$, $\eta^2_p = .21$ and a three-way interaction between block, incentive and sequence group, $F(1, 22) = 6.86$, $p = .016$, $\eta^2_p = .24$. The remaining effects were not significant, *F*s < 3.92, *p*s > .06.

In order to disentangle the three-way interaction between block, incentive and sequence group, separate 2 (incentive) x 2 (trial type) x 2 (sub-block) - ANOVAs for the two groups and the two blocks were computed. In the neutral start group, tonic pupil dilation did not differ between the first reward block and the first neutral block, $F(1, 12) = 2.99$, $p = .11$, but it was larger in the second reward block than in the second neutral block, $F(1, 12) = 22.85$, $p < .001$, $\eta^2_p = .66$. In the reward start group, this relationship was reversed: tonic pupil dilation was significantly larger in the first reward block than in the first neutral block, $F(1, 10) = 26.13$, $p < .001$, $\eta^2_p = .72$, but did not differ between the second reward block and the second neutral block, $F(1, 10) = 2.79$, $p = .13$ (see Figure 14).

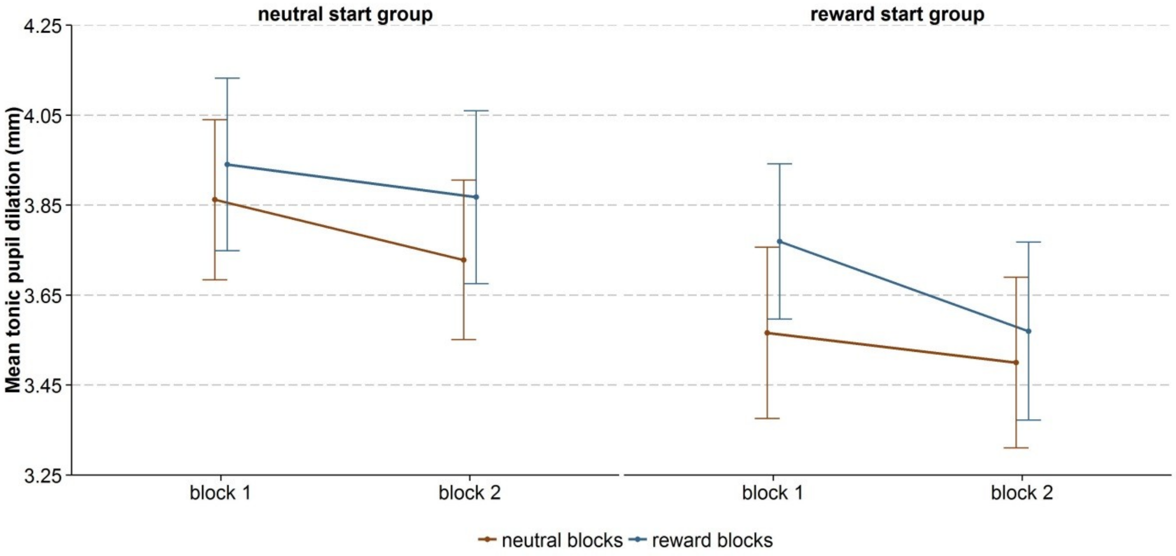

Figure 14. Mean tonic pupil dilation, separately for the two sequence groups, dependent on block and incentive. Error bars represent standard errors. Points are set off horizontally to enhance the visibility of the error bars (Study 3).

6.3.1.5 Phasic Pupil Dilation

Phasic pupil dilation was computed, like in Study 1, as the difference between the peak pupil diameter during the time between display onset and manual response and the baseline pupil diameter during the presentation of the fixation cross at the beginning of the trial. The 2 (incentive) x 2 (trial type) x 2 (block) x 2 (sub-block) x 2 (sequence group) - ANOVA on phasic pupil dilation yielded a main effect of incentive, $F(1, 22) = 7.81$, $p = .011$, $\eta^2_p = .26$ with larger pupil dilation in reward blocks. This main effect was qualified by two-way interactions with block, $F(1, 22) = 12.40$, $p = .002$, $\eta^2_p = .36$ and with trial type, $F(1, 22) = 5.75$, $p = .025$, $\eta^2_p = .21$. There were also a significant two-way interaction between block and sequence group, $F(1, 22) = 4.77$, $p = .04$, $\eta^2_p = .18$., a three-way interaction between block, sub-block and sequence group, $F(1, 22) = 7.56$, $p = .012$, $\eta^2_p = .26$, a four-way interaction between block, sub-block, incentive and sequence group, $F(1, 22) = 5.19$, $p = .033$, $\eta^2_p = .19$ and a five-way interaction between block, sub-block, incentive, trial type and sequence group, $F(1, 22) = 9.43$, $p = .006$, $\eta^2_p = .30$. The remaining effects were not significant, *F*s < 1.70, *p*s > .20.

In order to further examine the two-way Incentive x Block interaction, separate 2 (incentive) x 2 (trial type) x 2 (sub-block) x 2 (sequence group) - ANOVAs for the two blocks were computed. Phasic pupil dilation was larger in reward blocks compared to neutral blocks in the second blocks, $F(1, 22) = 15.34$, $p = .001$, $\eta^2_p = .41$, but not in the first blocks, $F < 1$ (see Figure 15). To disentangle the Incentive x Trial Type interaction, separate 2 (trial type) x 2 (block) x 2 (sub-block) x 2 (sequence group) – ANOVAs for reward and neutral blocks were conducted.

Phasic pupil dilation was not affected by trial type in the neutral blocks, $F(1, 22) = 1.57$, $p = .22$, but tended to be larger in CSD present trials than in CSD absent trials in the reward blocks, $F(1, 22) = 4.19$, $p = .053$, $\eta^2_p = .16$.

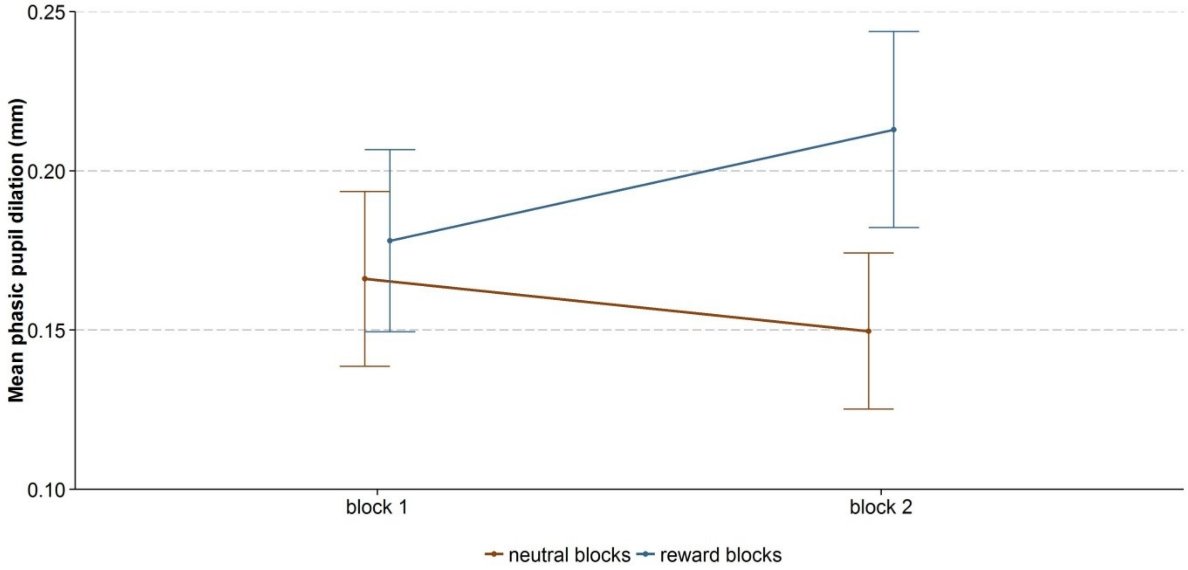

Figure 15. Mean phasic pupil dilation, dependent on incentive and block. Error bars represent standard errors. Points are set off horizontally to enhance the visibility of the error bars (Study 3).

6.3.1.6 Conclusion

Covert attentional capture, measured with response latencies, was reduced from block 1 to block 2 to a larger extent in reward blocks than in neutral blocks, but only because initial interference was higher in reward blocks. The level of interference in the second reward blocks was not significantly lower than the level of interference in the second neutral blocks. Thus, the results did not confirm Hypothesis 3.2 with regards to covert attentional capture.

On the other hand, reward increased mental effort both in the short-term (higher phasic pupil dilation in later reward blocks) and in the long-term (higher tonic pupil dilation in reward blocks), with the long-term effect being larger ($\eta^2_p = .61$ vs. $\eta^2_p = .26$). This confirms Hypothesis 3.1, but the effect of incentive interacted with block (phasic pupil dilation) and with block and sequence group (tonic pupil dilation).

For both tonic pupil dilation and response latencies, the three-way interaction between incentive, block and sequence group was significant. For accuracy, sequence group interacted with trial type and block. Participants in the reward start group did not show any interference effect of the CSD in their accuracy, whereas participants in the neutral start group did. These results suggest that trial type effects and reward effects depended at least partly on whether participants started the task with a neutral or with a reward block. Second, the aggregation of mean scores over the two neutral and the two reward blocks does not allow an analysis of practice effects across the four blocks, which were found in Study 2. To check whether the aggregation over two blocks may have blurred these practice effects, additional exploratory analyses with each of the four blocks as conditions of the variable block were conducted. This of course implied that incentive (neutral block vs. reward block) had to be excluded as an independent variable in this analysis.

6.3.2 Analysis 2: Exploratory Analysis of Practice Effects

6.3.2.1 Number of Fixations

On average, participants executed 2.75 fixations during a trial (95% CI [2.48, 3.01], range [1.71, 4.09]). The number of fixations made during a trial did not differ between the two sequence groups, $t(22) = 1.22$, $p = .24$ and was significantly higher than 2, $t(23) = 5.80$, $p <.001$, 95% CI [0.48, 1.01], indicating serial search. This important result demonstrates that response latency and accuracy cannot be analyzed as global indicators of attentional capture in this study. Nevertheless, the following analyses included both variables in order to detect possible reward effects on covert distractor interference. Furthermore, the oculomotor variables, such as the hit rate of the first saccade on the target, still can deliver information about the preattentive orienting of attention.

6.3.2.2 Accuracy

A 2 (trial type: CSD absent, CSD present) x 2 (sub-block: first, second) x 4 (block: 1 - 4) x 2 (sequence group: reward start, neutral start) - ANOVA on accuracy revealed a significant effect of trial type, $F(1, 22) = 6.42$, $p = .019$, $\eta^2_p = .23$, with more errors made in CSD present trials compared to CSD absent trials. This effect was modulated by sequence group, $F(1, 22) = 12.82$, $p = .002$, $\eta^2_p = .37$. Whereas participants in the neutral start group gave significantly more correct responses in CSD absent trials (M = 91%) than in CSD present trials (M = 86%), $t(12) = -5.44$, $p < .001$, 95% CI [.03, .07], participants' accuracy in the reward start group did not differ between the two trial types (Ms = 89%, 88%), $|t| < 1$. The interaction between block and sequence group was close to significance, $F(3, 66) = 2.69$, $p = .053$, $\eta^2_p = .11$. The remaining effects were not significant, Fs < 2.17, ps > .15.

6.3.2.3 Response Latencies

The 2 (trial type) x 2 (sub-block) x 4 (block) x 2 (sequence group) - ANOVA on response latencies yielded significant main effects of trial type, $F(1, 22) = 40.64$, $p < .001$, $\eta^2_p = .65$, sub-block, $F(1, 22) = 16.55$, $p = .001$, $\eta^2_p = .43$ and block, $F(3, 66) = 28.34$, $p < .001$, $\eta^2_p = .56$, with faster response latencies in the CSD absent condition, in the second sub-blocks and in later blocks. There was a significant interaction between block, trial type and sequence group, $F(3, 66) = 3.05$, $p = .035$, $\eta^2_p = .12$. Interference did not depend on sub-block, $F < 1$. All of the other effects were non-significant, Fs < 1.67, ps > .18.

To investigate the three-way Block x Trial Type x Sequence Group interaction, a 2 (trial type) x 2 (sub-block) x 4 (block) – ANOVA was run for each sequence group separately. In the neutral start group, the effects of trial type and block did not interact, $F < 1$. In the reward start group, the effect of trial type was modulated by the linear trend of block, $F(1, 10) = 16.59$, $p = .002$, $\eta^2_p = .62$. As can be seen in Figure 16, participants were more distracted by the CSD in the reward start group in the first block, but showed comparable performance to the other group in blocks 2 - 4. Mean interference in block 1 was marginally significantly larger in the reward start group, $t(22) = -2.07$, $p = .051$, but it did not significantly differ between the two groups in block 4, $t < 1$. To further examine practice effects on attentional capture, dependent t-tests comparing response latencies in CSD present and CSD absent trials were computed for each block in the two groups. These showed that the trial types did not differ in block 1 in the neutral start group, but in the later blocks. On the other hand, participants in the reward start group were distracted by the CSD in the first two blocks, but not later on (see Table 8).

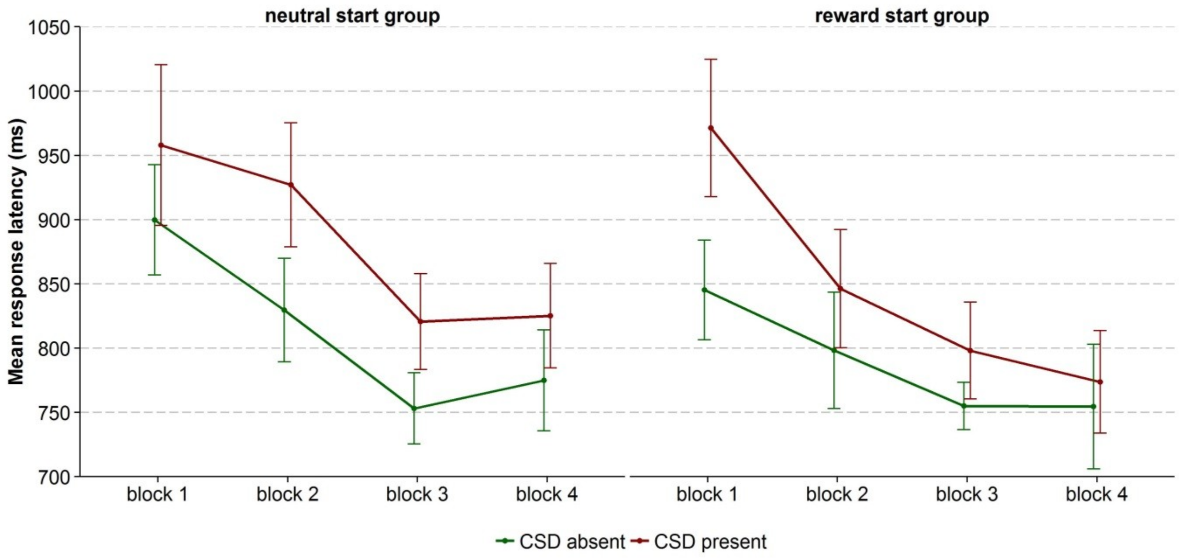

Figure 16. Mean response latencies, separately for the two sequence groups, dependent on block and trial type. Error bars represent standard errors. Points are set off horizontally to enhance the visibility of the error bars (Study 3).

Table 8. *Distractor Interference (ms) in the Four Blocks, Separately for the Two Sequence Groups (Study 3).*

		Distractor Interference			95% CI	
		M (SD)	*t(10)/t(12)*	*p*	*LL*	*UL*
Reward start group	Block 1	126 (82)	5.75	<.001	71	181
	Block 2	48 (58)	2.84	.018	9	87
	Block 3	43 (104)	1.22	.250	-26	113
	Block 4	19 (64)	1.58	.146	-23	62
Neutral start group	Block 1	58 (118)	1.76	.104	-13	130
	Block 2	98 (68)	5.27	<.001	57	138
	Block 3	68 (86)	2.72	.019	16	119
	Block 4	50 (65)	2.89	.013	11	89

Note. Distractor Interference = difference between mean response latency in CSD present trials and mean response latencies in CSD absent trials. LL = lower level of CI; UL = upper level of CI. *t*- and *p*-values are derived from *t*-Tests with the logarithmized response latencies; *M*, *SD* and 95% CI indicate the untransformed response latencies in ms.

6.3.2.4 Saccadic Latency

As the first saccade landed on the target only in 18% of the trials (see Chapter 6.3.2.5), too few participants had data in each cell of a 2 (trial type) x 2 (sub-block) x 4 (block) x 2 (sequence group) - ANOVA. Therefore, an analysis of this variable had to be skipped.

6.3.2.5 Location of First and Later Saccades

The hit rates of the first saccade and of all saccades on the different stimuli are presented in Table 9. Only 10 - 26% of the first saccades went to the target. This provides further evidence that search was more difficult and more serial in this task than in Study 1. Preattentive search for the target was especially unsuccessful in the neutral start group, where target hit rate was below or around the random probability (17%), and only increased a bit in the last block. The target hit rate of the first saccade seemed to be higher in the reward start group, especially in trials without the CSD. The hit rate of the first saccade on the CSD was also below the random hit rate in most conditions. The hit rate of all saccades on the CSD was markedly higher. The results for the form distractors were surprising because the hit rates both of the first and of all saccades were very low.

Table 9. *The Hit Rates of the First Saccade and of All Saccades on Different Stimuli (SD in Brackets), Separately for the two Sequence Groups, Dependent on Block and Trial Type (Study 3).*

		Neutral start group						Reward start group					
		First saccade			All saccades			First saccade			All saccades		
block		TA	FD	CSD	TA	FD	CSD	TA	FD	CSD	TA	FD	CSD
1	CSD absent	12 (9)	5 (4)	---	80 (26)	12 (4)	---	24 (13)	5 (3)	---	92 (12)	10 (4)	---
	CSD present	13 (7)	4 (3)	11 (9)	78 (25)	12 (6)	32 (15)	13 (7)	5 (4)	12 (9)	92 (7)	13 (4)	29 (14)
2	CSD absent	16 (9)	4 (4)	---	86 (12)	11 (6)	---	24 (14)	4 (3)	---	88 (13)	9 (5)	---
	CSD present	10 (8)	4 (3)	17 (10)	86 (12)	13 (7)	35 (15)	19 (14)	4 (2)	10 (7)	89 (14)	10 (4)	22 (16)
3	CSD absent	19 (17)	4 (3)	---	87 (12)	10 (5)	---	23 (17)	4 (3)	---	82 (29)	8 (4)	---
	CSD present	16 (10)	4 (3)	15 (15)	89 (9)	10 (6)	28 (17)	20 (11)	5 (3)	9 (9)	85 (24)	8 (5)	17 (12)
4	CSD absent	23 (19)	5 (4)	---	88 (15)	9 (6)	---	25 (16)	3 (3)	---	84 (23)	7 (4)	---
	CSD present	17 (9)	5 (5)	10 (10)	85 (15)	11 (7)	20 (12)	26 (17)	4 (4)	9 (7)	81 (22)	6 (6)	15 (14)

Note. TA = target, FD = form distractor, CSD = color singleton distractor.

6.3.2.6 Comparison of the Hit Rate on the CSD to the Hit Rate on the Form Distractor

The hit rate of the first saccade on the CSD was compared to that on the form distractors in CSD present trials with an ANOVA with saccade location (CSD, form distractor), block (1-4), sub-block (1-2) and sequence group as independent variables. This resulted in a main effect of saccade location, $F(1, 22) = 21.07$, $p < .001$, $\eta^2_p = .49$, with the hit rate on the CSD ($M = 12\%$) being larger than that on the form distractors ($M = 4\%$), The remaining effects were not significant, Fs < 2.71, ps > .11.

The same ANOVA was run on the hit rate of all saccades. Here, a main effect of the linear trend of block, $F(1, 22) = 51.61$, $p <.001$, $\eta^2_p = .70$ and sub-block, $F(1, 22) = 4.53$, $p = .045$, $\eta^2_p = .17$, emerged, with fewer saccades towards both distractor types in later blocks and in second sub-blocks. Importantly, the main effect of saccade location was significant again, $F(1, 22) = 47.56$, $p < .001$, $\eta^2_p = .68$ and interacted with the linear trend of block, $F(1, 22) = 18.15$, $p < .001$, $\eta^2_p = .45$. Whereas the hit rate of saccades towards form distractors remained stable across the four blocks, the proportion of CSD saccades declined linearly from block 1 to block 4 (see Table 9). All of the other effects did not reach significance, *F*s < 3.94, *p*s > .05.

6.3.2.7 Comparison of the Hit Rate of the First Saccade on the Target to the Hit Rate of the First Saccade on the CSD

The data for hit rates of the first saccade in the neutral start group suggest that the target and the CSD might both have drawn the same amount of attention, whereas the data for the incentive start group suggest that the target might have been more salient, at least in some conditions (see Table 9). To further investigate this, the hit rates of the first saccade on the target and on the CSD were analyzed in CSD present trials only. An ANOVA with saccade location (target, CSD), block (1 - 4), sub-block (1 - 2) and sequence group (reward start group, neutral start group) revealed a main effect of saccade location, $F(1, 22) = 5.96$, $p = .023$, $\eta^2_p = .21$, with more saccades being directed to the target (M = 17%) than to the CSD (M = 12%). This effect was modulated by an interaction with the linear trend of block, $F(1, 22) = 13.55$, $p = .001$, $\eta^2_p = .38$. The interaction between saccade location and group was significant, $F(1, 22) = 4.50$, $p = .045$, $\eta^2_p = .17$. The interaction between saccade location, group and block was marginally significant, $F(3, 66) = 2.28$, $p = .088$, $\eta^2_p = .09$. Furthermore, the linear trend of block was qualified by an interaction with sub-block and saccade location, $F(1, 22) = 6.85$, $p = .016$, $\eta^2_p = .24$. The remaining effects were not significant, *F*s < 3.38, *p*s > .07.

To disentangle the interaction between saccade location and group, the hit rates were compared in each group separately. In the neutral start group, the effect of saccade location was not significant, $F < 1$, but interacted with the quadratic trend of block, $F(1, 12) = 8.37$, $p = .013$, $\eta^2_p = .41$. This two-way interaction indicated that the CSD hit rate increased from block 1 to block 2 and then decreased until block 4, whereas the reverse pattern was observed for target hit rate. In the reward start group, the main effect of saccade location was significant $F(1, 10) = 10.89$, $p = .008$, $\eta^2_p = .52$, and interacted with the linear trend of block, $F(1, 10) = 9.59$, $p = .011$, $\eta^2_p = .49$. Participants showed a marked increase in their target hit rate from block 1 to block 4 (M = 13% to M = 26%), $t(10) = -3.23$, $p = .009$, 95% CI [-.23, -.04] whereas the CSD hit rate remained unchanged throughout the experiment. The difference in the percentage of target saccades in the fourth block between the two sequence groups was not significant, $t(22) = -1.72$, $p = .099$, 95% CI [-.21, .02] (see Figure 17).

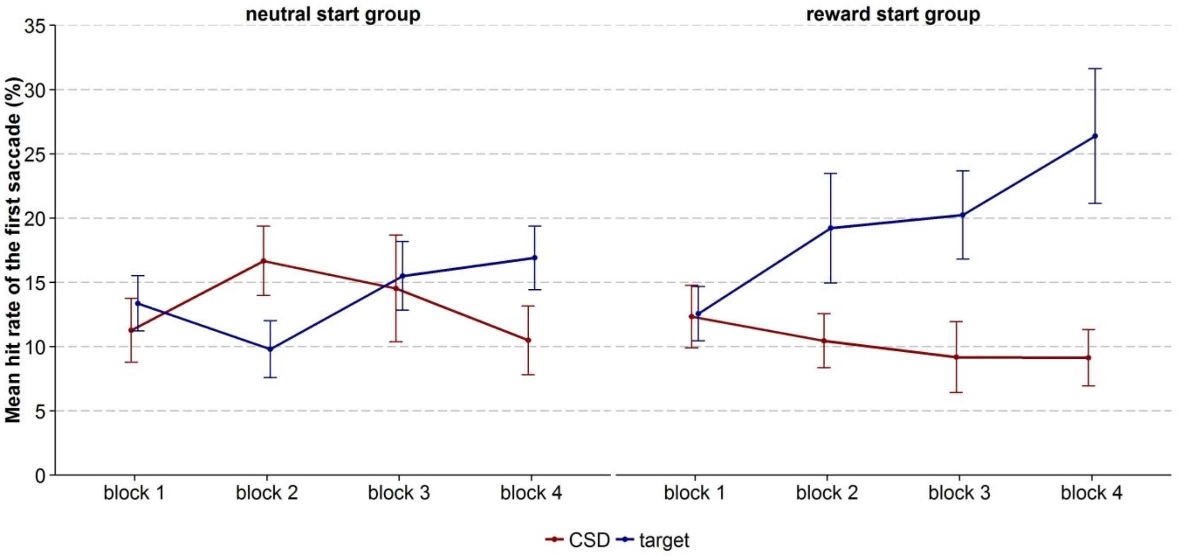

Figure 17. Mean hit rate of the first saccade on the CSD and on the target in CSD present trials, separately for the two sequence groups, dependent on block. Error bars represent standard errors. Points are set off horizontally to enhance the visibility of the error bars (Study 3).

6.3.2.8 *Hit rate of the First Saccade on the Target*

The previous chapter compared the hit rates of the first saccade on the target and on the CSD in CSD present trials. This chapter compares the hit rate of the target in CSD absent and CSD present trials in order to examine whether the CSD indirectly captured attention by reducing the hit rate on the target. The 2 (trial type) x 2 (sub-block) x 4 (block) x 2 (sequence group) - ANOVA on the hit rate of the first saccade on the target yielded a significant linear effect of block, $F(1, 22) = 21.22$, $p < .001$, $\eta^2_p = .49$ and trial type, $F(1, 22) = 6.26$ $p = .02$, $\eta^2_p = .22$, with more target saccades in later blocks and in CSD absent trials. The linear effect of block was modulated by trial type and sequence group, $F(1, 22) = 9.56$, $p = .005$, $\eta^2_p = .30$ and by sub-block and trial type, $F(1, 22) = 8.16$, $p = .009$, $\eta^2_p = .27$. The remaining effects did not reach significance, *F*s < 1.93, *ps* > .13.

To disentangle the three-way interaction between block, trial type and sequence group, separate ANOVAs for each sequence group were conducted. In the neutral start group, the interaction between the linear trend of block and trial type was not significant, $F(1, 12) = 2.90$, $p = .11$. In the reward start group, however, this two-way interaction was significant, $F(1, 10) = 6.23$, $p = .032$, $\eta^2_p = .38$. To further examine practice effects, dependent *t*-tests comparing target hit rates in CSD present and CSD absent trials were computed for each block in the two groups. In the reward start group, the CSD only interfered with hit rates in block 1, $t(10) = 3.87$, $p = .003$, but not later on, *ts* < 1.09, *ps* > .30. In the neutral start group, the distractor interfered with target saccades in the second block, $t(12) = 3.36$, $p = .006$, but not in the other blocks, *ts* < 1.73, *ps*

> 10. In block 2, participants in the reward start group made significantly more first saccades to the target than participants in the neutral start group (*M* = 20% vs. *M* = 15%), *F*(1, 22) = 4.83, *p* = .039. In the other blocks, both sequence groups did not differ in their amount of first target saccades, *F*s < 2.81, *ps* > .10 (see Figure 18).

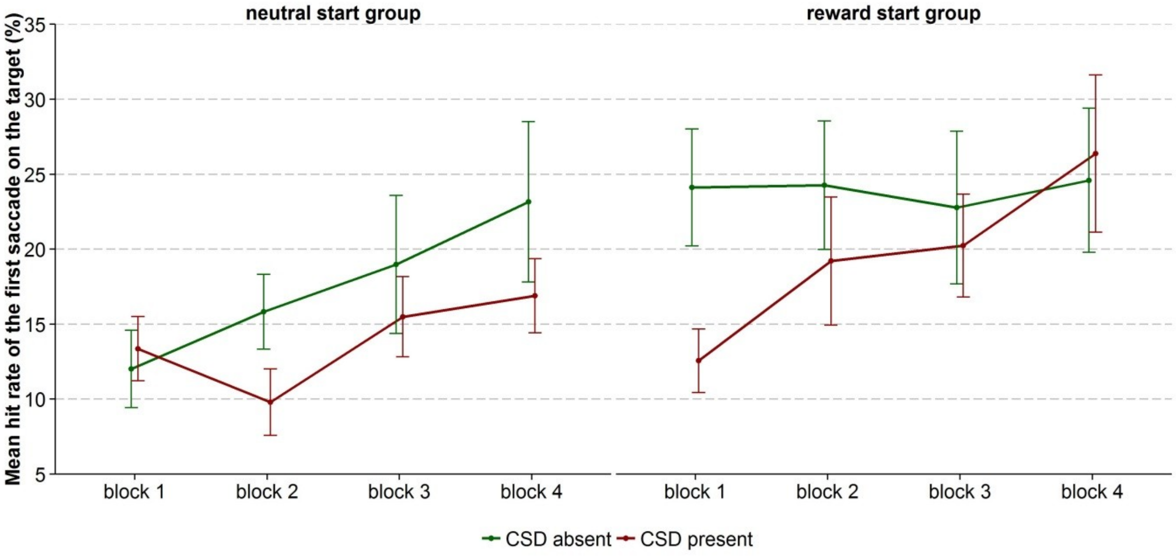

Figure 18. Mean hit rate of the first saccade on the target, separately for the two sequence groups, dependent on block and trial type. Error bars represent standard errors. Points are set off horizontally to enhance the visibility of the error bars (Study 3).

6.3.2.9 Fixation Duration on the CSD

To test the delayed disengagement account (Belopolsky et al., 2011; Fox et al., 2001; Fox et al., 2002), the mean fixation duration on the CSD was compared to the mean fixation duration on the form distractors in an ANOVA with fixation location (CSD, form distractor), block (1-4), sub-block (1-2) and sequence group as independent variables. The ANOVA revealed a significant main effect for the linear trend of block, *F*(1, 22) = 18.43, *p* < .001, η^2_p = .46, with shorter fixations in later blocks, and an interaction between fixation location and sequence group, *F*(1, 22) = 7.53, *p* = .012, η^2_p = .26. The remaining effects were non-significant, *F*s < 3.04, *ps* > .09. In the neutral start group, the effect of fixation location was not significant, *F*(1, 12) = 1.19, *p* = .30, but participants in the reward start group fixated longer on the form distractors (*M* = 123 ms) than on the CSD (*M* = 95 ms), *F*(1, 10) = 7.03, *p* = .024, η^2_p = .41.

6.3.2.10 Fixation Duration on the Target

The 2 (trial type) x 2 (sub-block) x 4 (block) x 2 (sequence group) - ANOVA on the fixation duration on the target revealed a main effect of the linear trend of block, *F*(1, 22) = 13.84, *p* = .001, η^2_p = .39, with shorter target fixations in later blocks and a main effect of sequence group, $F(1, 22) = 4.89$, $p = .038$, $\eta^2_p = .18$, with shorter target fixations in the neutral start group (M = 148 ms) than in the reward start group (M = 168 ms). The remaining effects were not significant, Fs < 2.00, ps > .12.

6.3.2.11 Tonic Pupil Dilation

The 2 (trial type) x 2 (sub-block) x 4 (block) x 2 (sequence group) - ANOVA on tonic pupil dilation yielded significant main effects of block, $F(3, 66) = 14.30$, $p < .001$, $\eta^2_p = .39$ and sub-block, $F(1, 22) = 59.40$, $p < .001$, $\eta^2_p = .73$, with larger pupil diameters in the first blocks and sub-blocks. Sequence group interacted with the linear, $F(1, 22) = 9.35$, $p = .006$, $\eta^2_p = .30$ and with the cubic trend of block, $F(1, 22) = 34.07$, $p<.001$, $\eta^2_p = .61$. There was a significant Block x Sub-Block x Sequence Group interaction, $F(3, 66) = 2.77$, $p = .049$, $\eta^2_p = .11$. The remaining effects were not significant, Fs < 2.52, ps > .06.

In the neutral start group, the cubic trend of block was significant, $F(1, 12) = 32.36$, $p < .001$, $\eta^2_p = .73$. In the reward start group, the linear trend of block was significant, $F(1, 10) = 21.20$, $p= .001$, $\eta^2_p = .70$. In both groups, tonic pupil diameter dropped after the first reward block (block 1 in the reward start group and block 2 in the neutral start group) (see Figure 19). This trend-analysis explains the pattern of results of Analysis 1 for tonic pupil dilation (see Chapter 6.3.1.4). When mean tonic pupil dilation is collapsed across neutral and reward blocks, participants in the neutral start group show a reward effect only in the later blocks (the difference between block 3 and block 4 in Figure 19), whereas participants in the reward start group only show this effect in the first blocks (the difference between block 1 and block 2 in Figure 19).

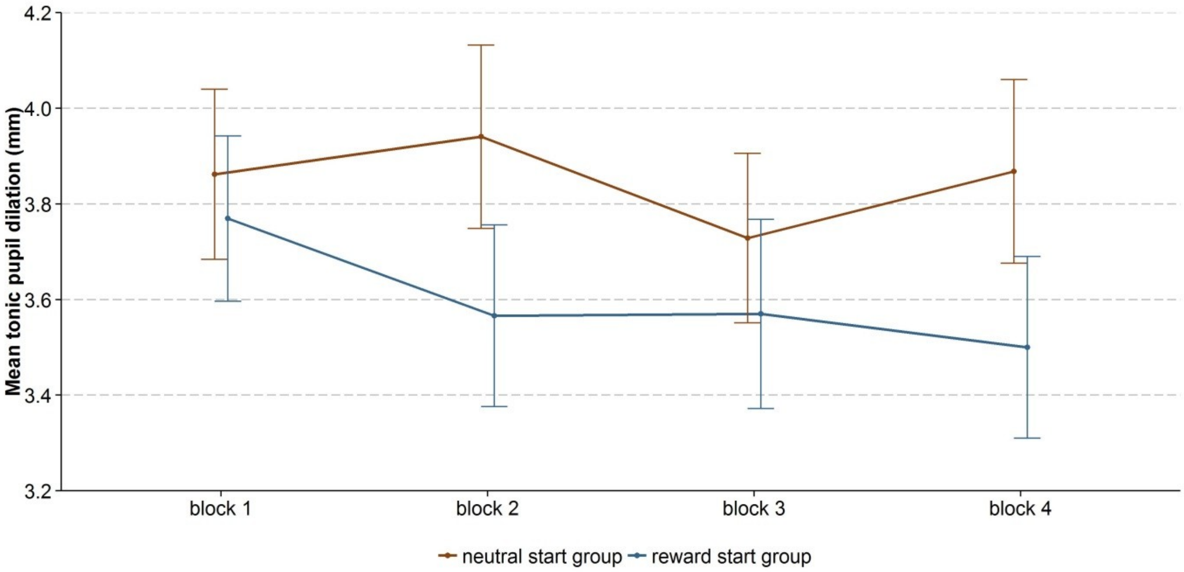

Figure 19. Mean tonic pupil dilation, separately for the two sequence groups, dependent on block. Error bars represent standard errors. Points are set off horizontally to enhance the visibility of the error bars (Study 3).

6.3.2.12 Phasic Pupil Dilation

The 2 (trial type) x 2 (sub-block) x 4 (block) x 2 (sequence group) - ANOVA on phasic pupil dilation revealed a significant two-way interaction between the linear trend of block and sequence group, $F(1, 22) = 12.50$, $p = .002$, $\eta^2_p = .36$, between the quadratic trend of block and sequence group, $F(1, 22) = 12.40$, $p = .002$, $\eta^2_p = .36$ and between the cubic trend of block and sequence group, $F(1, 22) = 4.95$, $p = .037$, $\eta^2_p = .18$. The remaining effects were not significant, *F*s < 2.24, *p*s > .08.

In the neutral start group, only the linear effect of block was significant, $F(1, 12) = 11.64$, $p = .005$, $\eta^2_p = .49$. In the reward start group, the quadratic, $F(1, 10) = 8.49$, $p = .015$, $\eta^2_p = .46$, and the cubic trend, $F(1, 10) = 5.51$, $p = .041$, $\eta^2_p = .36$, were significant. Interestingly, pupil dilation only increased in both groups in the second reward block (block 3 in the reward start group and block 4 in the neutral start group) compared to the previous block (see Figure 20). This explains the finding of Analysis 1 that phasic pupil dilation was increased only in the second reward blocks (see Chapter 6.3.1.5).

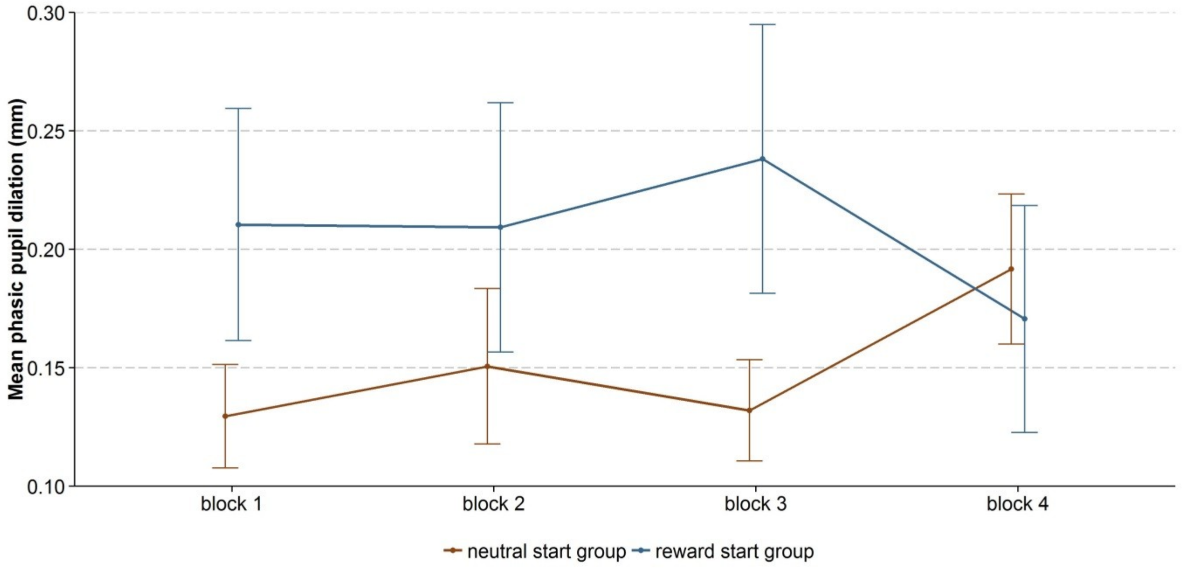

Figure 20 Mean phasic pupil dilation, separately for the two sequence groups, dependent on block. Error bars represent standard errors. Points are set off horizontally to enhance the visibility of the error bars (Study 3).

6.3.2.13 Correlation Between Response Latencies and Phasic Pupil Dilation

To test if the increase in phasic pupil dilation in incentive blocks correlated with a decrease of response latencies in the same blocks, difference scores for the two variables were computed. These were based on the trend analysis for phasic pupil dilation, which showed a marked increase from the second to the third and from the third to the fourth block, respectively. In the reward start group, the difference between the mean values in the third and in the second block was computed for both variables. In the neutral start group, the mean values in the third block were subtracted from those in the fourth. In both cases, positive values indicate an increase. Thus, a negative correlation between the two difference scores would be expected if an increase in mental effort was associated with a decrease in response latencies. As the decrease in response latencies was modulated by sequence group, the correlation was computed for each sequence group separately. Analysis showed that the variables neither correlated in the reward start group, $r(11) = .09$, $p = .79$, nor in the neutral start group, $r(13) = .38$, $p = .20$. Furthermore, the correlation was clearly not negative.

6.4 Discussion

Study 3 tested three hypotheses. First, it was assumed that performance-contingent reward increases mental effort as indicated by an increase in phasic (short-term effects) and tonic (long-term effects) pupil dilation in incentive blocks compared to neutral blocks (Hypothesis 3.1). Second, it was expected that reward increases the search performance

(Hypothesis 3.2) and third, it was hypothesized that attentional capture is decreased in reward blocks, compared to neutral blocks (Hypothesis 3.3). The number of fixations executed during a trial indicated that the task which was employed required serial search. In contrast to Study 1, the average number of fixations was significantly higher than two. The analysis of the hit rate of the first saccade on the target confirmed that search was definitely more difficult than in Study 1. This hit rate was much lower in Study 3 than in Study 1 (M = 19% vs. M = 37%), while the overall target hit rate did not differ that much (M = 78% vs. M = 86%). On the other hand, mean fixation duration on the target in Study 3 (158 ms) was similar to the one in Study 1 (151 ms). This is an interesting finding given that the target was variable in the former. Study 3 employed the more difficult version of the additional singleton task in which the target and the distractor changed their form randomly (cf. Theeuwes, 1991). This was done because the results of Study 1 and Study 2 indicated that attentional capture in the additional singleton task is volatile and quickly vanishes with sufficient practice. Therefore, attentional capture only occurs very markedly during a limited number of trials during the experiment. This posed the question of how such a “short-lived” but strong interference could be modulated by reward manipulations. Consequently, the more difficult variant of the additional singleton task was employed in Study 3 because it typically causes a greater amount of attentional capture which was assumed to be present during a higher number of trials.

Theeuwes (1991) and, more recently, Wentura et al. (2013) reported flat search slopes with this kind of task. The decisive difference between Study 3 and those studies might be that the latter did not measure eye movements and presented the stimuli very close to the fixation center. The stimulus eccentricity in the study by Theeuwes (1991) was 3.4°. Wentura et al. (2013) did not report target eccentricity. Response latency and accuracy could not be used as indicators of attentional capture in the present study because they reflect both the preattentive and the serial process of search. The hit rate of the first saccade, however, is solely determined by the first preattentive orienting of attention. Consequently, although search might not have been completely parallel in this task, the analysis of this variable still allowed making inferences about reward effects on attentional capture. Analyses with the indicators of the general search process (response latencies, accuracy, and overt indicators of attentional orienting) were also conducted to find out whether distractor interference in general was modulated by incentives. The first two chapters of this discussion summarize and interpret the results of Study 3 with reference to their meaning for Hypothesis 3.1 (see Chapter 6.4.1) and Hypothesis 3.2 and Hypothesis 3.3 (see Chapter 6.4.2). The third chapter (6.4.3) addresses some limitations of Study 3 and the last chapter (6.4.4) draws a conclusion.

6.4.1 Incentives Increase Transient and Sustained Mental Effort

Hypothesis 3.1 was partly confirmed. Analysis 1 showed significant main effects of incentive both for tonic and phasic pupil dilation, yet these were modulated by block and sequence group (tonic pupil dilation) or block (phasic pupil dilation). Analysis 2 revealed that incentives increased transient and sustained mental effort only in the second incentive block. This suggests that participants first had to experience the effect of reward during the first incentive block before they intentionally invested more effort during the second reward block. As was expected, the transient effort recruitment in neutral blocks remained stable. The present results are consistent with previous studies, which found a main effect of incentive on phasic pupil dilation (Bijleveld et al., 2009; Chiew & Braver, 2013). Those studies did not report an analysis of block or sub-block effects on the incentive effect. Bijleveld et al. (2009) employed only 48 trials. Chiew and Braver (2013) used only one incentive and one neutral block, with 200 trials each. This, however, does not preclude the possibility that practice effects occurred within the block. The analysis of Study 3 suggests that incentives in a difficult search task with variable targets first have to be experienced before they lead to an increase in mental effort.

The interpretation of the results for tonic pupil dilation is more complex. Analysis 1 showed that it declined from the first to the second blocks and that this trend was modulated by sequence group and incentive. Analysis 2 demonstrated that the general decline in tonic pupil dilation over the four blocks was solely due to the participants in the incentive start group, whereas the pupil diameter in the neutral start group showed a cubic trend. Thus, the incentive effect occurred only at the start of the experiment in the reward start group (decline from the first to the second block), but in the middle (decline from the second to the third block) and at the end of the experiment (increase from the third to the fourth block) in the neutral start group. The results in the neutral start group are close to what was expected to be found in the whole sample: an increase of tonic pupil dilation in an incentive block, followed by a decrease in a neutral block. This finding is consistent with research demonstrating that incentives that are constantly given during a longer period of time also increase tonic pupil dilation (Chiew & Braver, 2013; Heitz et al., 2008). The results in the reward start group, however, are more difficult to interpret. As tonic pupil dilation remained stable across block 2 to block 4 in this group, it is questionable if the decrease from block 1 to block 2 can be interpreted as an incentive effect or if it reflects a general practice effect.

Finally, tonic pupil dilation tended to be higher in the neutral start group and phasic pupil dilation tended to be higher in the reward start group. Although these differences were not significant in any of the analyses, it would be interesting to follow up on this issue. To arrive at a valid differentiation and interpretation of tonic and phasic pupil dilation, more research on their association with performance indicators has to be carried out. In this Study, like in Study 1, the difference scores of phasic pupil dilation and response latency did not correlate, which could be due to the low sample size in the sequence groups. It is important to note that the study by Wykowska et al. (2013) is the only study so far that has observed a significant correlation between the change in (tonic) pupil dilation and an performance indicator during a visual search task. The other studies did not report whether they have tested for correlation between pupil dilation and performance (Bijleveld et al., 2009; Chiew & Braver, 2013; Heitz et al., 2008). Another reason for the absence of a correlation between pupil dilation and response latency in the present study might be that pupil dilation does not only measure mental effort, but also emotional arousal. This issue is further addressed in the discussion in Chapter 8.

6.4.2 Attentional Capture and Distractor Interference Depend on the First Experience with the Color Singleton Distractor

The search process in Study 3 was not only different from the one in Study 2 because search was serial in the former, but also because interference was not modulated by the variable sub-block. This result suggests that specific distractor suppression depends on a constant target, as it was used in Study 2 and in the study by Vatterott and Vecera (2012). With the variable target used in Study 3, however, distractor interference was stable within each of the blocks. Thus, a trial-wise manipulation of incentives (cf. Study 1) would have been adequate for this task. Based on the results of Study 2, reward was manipulated block-wise in the present study because interference was thought to be only present in the first half of a block (see Chapter 5.4.4).

Analysis 1 (see Chapter 6.3.1) collapsed dependent variables across the two blocks of the same incentive condition and did not confirm Hypotheses 3.2 and 3.3. Firstly, search performance, as indicated by response latencies, accuracy or hit rates of the first saccade, was not improved in reward blocks compared to neutral blocks. Consequently, the results did not support Hypothesis 3.2. Secondly, interference in response latency decreased in reward blocks relative to neutral blocks, yet only because interference was larger in the first reward block. This was surprising because mental effort was clearly increased in reward blocks. Figure 17 and Figure 18, displaying the hit rate of the first saccade on the target and on the CSD in Analysis 2, show that attentional capture was not reduced in reward blocks compared to incentive blocks, either. Thus, Hypothesis 3.3 was not confirmed.

Importantly, both pupil and manual response data showed that participants' response varied considerably depending on the sequence group they were in. The two sequence groups did not differ in their amount of reward or in any of the dependent variables, when these were collapsed across conditions, but sequence group almost always interacted with incentive and other independent variables. The sequence of reward and neutral blocks had a moderating effect on the role of incentives.

Because of this reason, and because Study 2 provided partial support that distractor interference does not only vary between sub-blocks, but also between blocks, a second analysis was conducted that differentiated between each of the four blocks (see Chapter 6.3.2). This analysis, however, was exploratory because hypotheses were based on the assumption that reward effects do not depend on the sequence of reward and neutral blocks. Thus, it is important to point out that the design of Study 3 was not fully suitable to study group differences because only eleven, respectively 13 participants were in each sequence group. This leaves the possibility that any group differences might be a random finding and disappear with larger sample sizes. Due to the large number of dependent variables, the results of Analysis 2 are summarized for the most important variables separately.

Indicators of Covert Distractor Interference: Response Latencies and Accuracy

The results showed that covert distractor interference was marginally significantly larger in the reward start group than in the neutral start group in the first block. From block 3 on, interference was absent in the reward start group. In the neutral start group, distractor interference was only present from block 2 to block 4. These results showed that Analysis 1 discarded important information: Distractor interference was increased in reward trials compared to neutral trials, but this difference was solely due to the first block in the reward start group. The fact that participants in the reward start group were not distracted by the CSD in block 3 and block 4 suggests a facilitation effect of a rewarding start on general distractor suppression. In addition, it is rather surprising that participants in the neutral start group were not distracted by the CSD in the first block, but only later on. On the other hand, the tendency of responses to be slower in the CSD present condition was clearly present (p = .10) and, more important, the CSD was fixated significantly more often than neutral distractors, especially in the beginning of the experiment. Furthermore, results for accuracy suggest that interference, in contrast to Study 2, was not only observed in response latencies: Participants in the neutral start group made significantly more errors in CSD present trials than in CSD absent trials, whereas individuals' accuracy in the other group did not differ between the trial types, and they did not make significantly more errors than the other group. Thus, participants in the neutral start group made more errors in CSD present trials, without being significantly faster (there was no interaction between trial type and sequence group for response latencies). Convergent evidence comes from the data on the fixation duration on the target: the target fixations in the neutral start group were significantly briefer than the ones in the reward start group. In general, mean accuracy was much higher than in Study 1 (Study 3: M = 89%; Study 1: M = 81%). This indicates that the punishment of errors, which was introduced in Study 3, had a beneficial effect on overall accuracy.

Indicators of Overt Attentional Capture: Hit rate of the First Saccade on the Target and on the CSD.

In contrast to response latencies and accuracy, the hit rate of the first saccade is an indicator of attentional capture. First, the CSD itself was hit by first saccades more often than form distractors. Second, the CSD indirectly captured attention by reducing the percentage of target saccades.

The hit rate on the CSD (M = 12 %) was lower than that on the target (M = 17%), but higher than that on the form distractors, which was very low (M = 4%). Although the hit rate on the CSD was well below the random hit rate, this result still has a meaning for attentional capture because participants were able to almost completely suppress the form distractors, but not the CSD. This suggests that they did not randomly fixate any of the six stimuli locations; for this to be true the hit rate on form distractors would need to be at least 17%. Compared to Theeuwes et al. (2003), who reported that the first saccade landed on the CSD in 38% of the trials, the hit rate of the first saccade on the CSD was quite low in Study 3, probably because search was serial (the mean hit rate of all saccades on the CSD ranged from 15% to 35% in the different conditions, suggesting that the CSD mainly affected the later orienting of attention). This direct form of attentional capture was not modulated by the sequence group and remained stable across the blocks.

On the other hand, the hit rate of the first saccade on the target was affected by the sequence group. A direct comparison of target and CSD saccades for CSD trials only showed that participants in the reward start group were faster in making their search for the target more efficient. In both sequence groups, search for the target was very inefficient in the first block, in which the hit rate on the target was equal to that on the CSD. However, participants in the reward start group increased their target hit rate faster than those in the neutral start group. Furthermore, participants in the reward start group did not show an increase in their hit rate on the CSD, whereas the hit rate on the CSD increased from the first to the second block in the neutral start group. Similar results were revealed by the analysis of the target hit rate in both CSD absent and CSD present trials: Participants in the reward start group were faster at reducing attentional capture because they increased their target hit rate in CSD present trials while keeping constant the proportion of target saccades on CSD absent trials.

Overt Distractor Interference: Hit Rate of All Saccades on the CSD

The hit rate of all saccades on the CSD was significantly higher than the hit rate on the form distractors. Unlike the hit rate of the first saccade, the hit rate of all saccades on the CSD declined over the course of the experiment. Participants in both groups learned to ignore the distractor with their gaze. However, this practice effect was associated with decreased interference in response latencies only in the reward start group. In the reward start group, participants' manual response was not delayed by the CSD in block 3 and block 4, which could be explained with the decrease of CSD saccades. In the neutral start group however, participants' response latency showed a CSD effect in the last block despite the reduced rate of CSD saccades. This is indeed a peculiar finding that calls for further research.

Fixation Duration on the Target and on the CSD

The fixation duration data cannot explain why the reduction of oculomotor interference by the CSD was not associated with a reduction in response latency interference in the neutral start group. The results also speak against an attentional-disengagement account of attentional capture. In the neutral start group, fixation durations on the CSD did not differ from those on the form distractors and in the reward start group, participants even fixated longer on the form distractors than on the CSD. Thus, it can be concluded that the CSD interfered with the execution of target saccades and was fixated quite often during the serial stage of search, but was quickly recognized as an irrelevant distractor and did not hold attention any longer than the form distractors.

Summary

As has been mentioned, Hypotheses 3.2 and 3.3 were not confirmed by the results. Search performance was not increased and distractor interference was not reduced in reward blocks compared to neutral blocks, but the experience in the first block was decisive: A rewarding start facilitated general distractor suppression as indicated by a) less interference in accuracy during the whole experiment (see Chapter 6.3.2.2), b) less interference in response latencies in later blocks (see Chapter 6.3.2.3), c) a higher hit rate of the first saccade on the target, compared to the CSD (see Chapter 6.3.2.7) and d) an earlier and more marked decrease of interference in the hit rate of the first saccade on the target (see Chapter 6.3.2.8). However, it is important to note that the two sequence groups did not differ in the absolute level of their performance, measured with these variables (i.e., mean response latency, mean accuracy and mean hit rate on the target) in the last block of the experiment.

A possible explanation for the sequence effect of reward argues that participants not only have to learn to suppress the CSD, but also have to learn the procedure and the rules of the payoff scheme. Concerning the latter, participants have to experience that they cannot win a reward in every trial of a block because of the adaptive time limit. In the beginning, participants might commit themselves to the goal to win in all of the trials because they are not aware of the fact that the limit decreases with each decrease in response latency. Thus, the learning process was maybe "smoother" in the reward start group because participants in this group completed both learning processes in the first block: they established an effective distractor suppression strategy and learned how much money they could win over the course of the experiment. In the neutral start group, however, these processes were distributed among the first and the second block. Obviously, if one combines the first experience with the CSD with the first experience with the payoff strategy, participants show the best performance in the long run. The reward learning process might be accelerated by explicitly informing participants about the adaptive response time limit in advance.

6.4.3 Limitations

As has been pointed out, the most serious limitation of this study is that it was not designed to conduct between-subjects comparisons, and thus all of the results of Analysis 2 are exploratory. Furthermore, the sample sizes in the two groups are quite small (11 vs. 13 participants). Additional studies with larger sample sizes are needed to test whether the results of Analysis 2, especially those concerning the effects of sequence group, are replicable.

A fixation check procedure was added to the experiment to guarantee valid fixations throughout the trial. This was successful in reducing the amount of trials that could not be included in analysis because eye movements were not recorded: These made up 7.4% of the trials in Study 1, in which the fixation check procedure was not implemented, but only 1.8% of the trials in Study 3. Yet, in comparison to Study 1, the amount of anticipatory saccades increased in Study 3 from 4.9% to 7.4%. Theoretically, anticipatory saccades could be detected by the experimental program by breaking the trial procedure if an early first saccade is registered. Such a procedure, however, was not implemented in any of the studies cited in this thesis and it would require thorough pre-testing of the detection mechanism so that it does not detect too many false positives. Like in Study 1, an analysis was run that included every trial with a correct manual response in order to check whether the exclusion of trials led to a bias in the data. This yielded the same pattern of results for response latencies and accuracy.

The results for phasic pupil dilation - an undisputed indicator of mental effort - showed that it was increased, as expected, in reward blocks compared to neutral blocks. However, this increase only occurred in the comparison between the second reward block and the second neutral block. This might be a further hint that participants needed at least one block to a) learn the specific task and b) learn the payoff scheme in order to increase their task engagement in the second half of the experiment. Thus, incentives might not have an immediate effect on mental effort, but only after participants have learned how they are related to their task performance and how large in size they are. Therefore, it is important to consider this in the task procedure and to employ a sufficient number of trials. This experiment contained far less incentive trials than Study 1 (288 in Study 1, 96 in Study 3) and less practice trials (60 in Study 1, 20 in Study 3), which might be a reason why the hypothesized reward effects were not evident, especially given that the task in Study 3 was more difficult. On the other hand, Study 3 contained almost the same number of incentive trials as the study by Chiew and Braver (2013). This suggests either that top-down control in Study 3 was more difficult or that the payoff scheme in Study 3 was less motivating.

6.4.4 Conclusion

This study could only contribute limited evidence to answer the question whether motivation enhances top-down control over attentional capture because of two reasons: Firstly, the task was not easy enough to be performed with parallel search. By solving one problem (using a paradigm in which interference is so large that it does not vanish after some trials), another problem arose (search is so difficult that it cannot be executed in parallel). The issue of trying to measure attentional capture with eye movements is discussed in more detail in Chapter 8. Therefore, only two variables (the hit rate of the first saccade on the CSD and on the target) could be analyzed as indicators of attentional capture. Strong interference by the CSD was observed in response latencies and in accuracy, yet as search needed significantly more saccades than Study 1 and the average number of fixations was significantly higher than two, these variables cannot be interpreted as indicators of attentional capture because they consist of both preattentive and serial processes. Secondly, Analysis 1 disconfirmed Hypothesis 3.3, which expected a decrease of attentional capture in reward blocks compared to neutral blocks, whereas Analysis 2 showed that it does make a difference whether an observer starts the search task earning rewards or not.

Another interesting insight gained by Study 3 is that experience-specific distractor suppression (Vatterott & Vecera, 2012) is dependent on a constant target: Unlike in Study 2, attentional capture or distractor interference was not modulated by sub-block. Thus, this version of the additional singleton task is adequate for a trial-wise manipulation of incentives, which circumvents possible confounds between reward effects and sequence effects in a design with blocked rewards.

7 STUDY 4: EMOTIONAL FACES DO NOT CAPTURE ATTENTION, BUT INTERFERE WITH SERIAL SEARCH

7.1 Hypotheses

The hypotheses of Study 4 were derived from the theoretical background outlined in Chapter 2.3:

- **Hypothesis 4.1:** An emotional face does not capture covert and overt attention if its emotional expression is irrelevant for the task goal (and if search is parallel).
- **Hypothesis 4.2:** An emotional face interferes with serial visual search, either by delaying disengagement of attention or by being fixated more often than faces without emotional expression.
- **Hypothesis 4.3:** The distraction effect described in Hypothesis 4.2 is larger for happy faces than for angry faces.

Two different search tasks were compared in a between-subjects design in order to investigate these hypotheses. The task goal for one group of participants was identical to the one used in the experiments by Hodsoll et al. (2011): Participants had to search for the male face among female faces. This task was assumed to induce serial search and emotion should only distract attention during focal processing. The other group searched for a target whose attributes were independent of emotional content and which could be found with parallel search. One possibility would have been to use the same target as the study by Huang et al. (2011): a black dot on one of the faces. Yet, this could lead to visual confounds between the black dot and the different photographic face stimuli. Thus, a procedure similar to the study by Devue, Belopolsky, and Theeuwes (2012) was used. Next to the face stimuli, colored dots were positioned and participants were instructed to search for the differently colored dot. This is a classic pop-out search task that allows for parallel search. Like in the experiments by Hunt et al. (2007) and by Devue et al. (2012), participants' task was to make a saccade to the target face as quickly as possible. Participants did not give a manual response to the target. Hypothesis 4.3 is based on the results by Becker et al. (2011), who controlled for visual confounds and observed a happiness-superiority effect.

7.2 Method

7.2.1 Participants

41 students (28 female) of the University of Heidelberg took part in the experiment. Their age ranged from 18 to 45 years (*M* = 22 years). The participants received course credit as compensation.

7.2.2 Design

Trial type and *face orientation* were varied within participants. The display in a trial could either consist of six neutral stimuli (*all neutral*), or could contain a *happy target*, an *angry target*, a *happy distractor* or an *angry distractor* (cf. Huang et al., 2011). In the latter four trial types, only one face contained an emotional expression, so that either the target or one of the distractors carried emotional content. Face orientation was manipulated block-wise, so that participants searched among upright faces in two blocks and among inverted faces in the remaining two blocks. The sequence of blocks was counterbalanced among participants: one group started with two blocks with inverted faces and continued with two blocks of upright faces and the other group started with upright faces and continued with inverted faces. Each trial type was repeated 12 times in a block, thus each combination of trial type and face orientation was repeated 24 times throughout the experiment. The *task goal* was varied between participants. Participants in the *male target group* (MTG) had to make a saccade to the singleton male face among female faces (cf. Hodsoll et al., 2011). Participants in the *dot target group* (DTG) searched among a different display: face gender of the target was factorially combined with trial type and face orientation, and the gender of the remaining distractor faces was randomly determined. Next to each face, an orange or pink dot was positioned. All but one dot shared the same color. Participants' task was to make a saccade to the face next to the dot with the unique color. Figure 21 shows examples for the display in the MTG and in the DTG with inverted and upright faces.

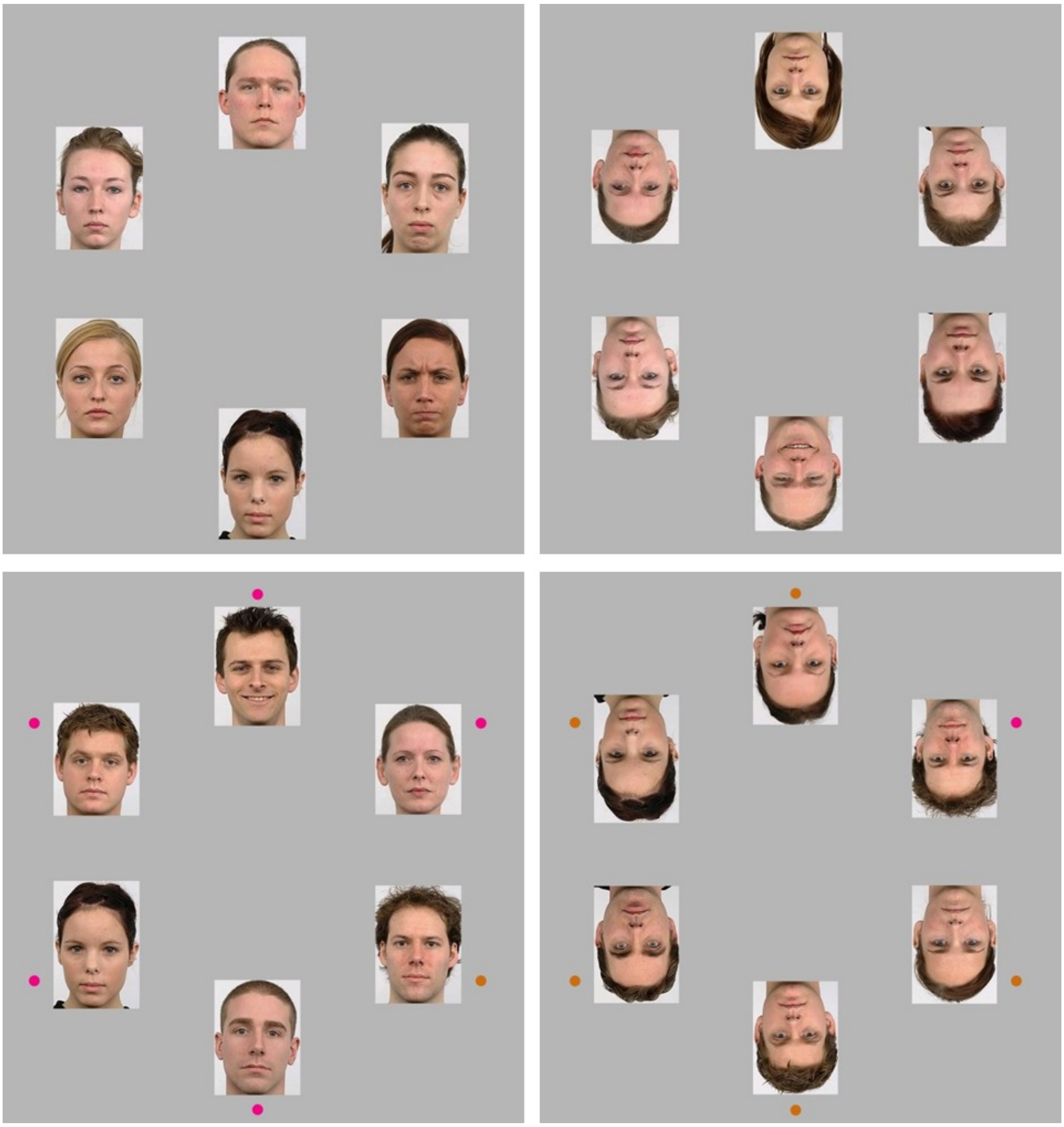

Figure 21. Examples of different display types in the two task goal groups. In the first row, an angry distractor trial (left) and a happy target trial (right) in the MTG with upright and inverted faces. In the second row, a happy distractor trial (left) and an all neutral trial (right) in the DTG with upright and inverted faces (Study 4).

7.2.3 Stimuli

The face stimuli were taken from the Radboud Faces Database (Langner et al., 2010), which contains 67 models displaying eight emotions. This database was chosen because the facial expressions are more valid (mean agreement rate between the expressed emotion and the perceived emotion by test participants: 82%) than those of the Karolinska Directed Emotional Faces (mean agreement rate: 72%), the most popular and one of the largest facial databases

(Goeleven, De Raedt, Leyman, & Verschuere, 2008; Langner et al., 2010). Only Caucasian models were selected to avoid confounds due to the own-race bias (Hills & Pake, 2013). Among those, the neutral, the happy and the angry expression of all the 19 female adult and 19 of the 20 male adult models (i.e., all of the models except for number 71) with frontal gaze direction were used. The pictures were 150 pixels wide and 200 pixels high (5.1° x 6.9° on screen). In contrast to the study by Hodsoll et al. (2011), hair was not cropped to increase ecological validity. The dots were circles with a diameter of 0.5° and were either orange (RGB: 202, 110, 18) or pink (RGB: 242, 0, 135). The six face stimuli were positioned on an imaginary circle with a radius of 8.6°. In the DTG, the colored dots were arranged on an imaginary circle with a radius of 11.9°.

7.2.4 Procedure

After participants had signed an informed consent and had read the instructions, the eye tracker was calibrated. 13 small squares appeared in different positions on the screen one after another and had to be fixated by the observer. In case of a successful calibration, the experimental task started. Each trial began with the presentation of the central black (RGB: 0, 0, 0) fixation cross (0.5° x 0.5°) on grey (RGB: 180, 180, 180) background. If the recorded gaze position fell within a radius of 50 pixels around the center of the fixation cross, and the fixation cross was already presented on screen for at least 500 ms (in order to control for anticipatory fixations), the fixation was regarded as valid. In case of a valid fixation, the fixation cross remained on screen for another 500 ms before the search display was presented. If a valid fixation did not fall on the fixation cross within 2000 ms after its initial presentation, a recalibration was initiated. After the fixation cross disappeared, the six face stimuli (and, in the DTG, the six dots) were presented on the grey background. Participants had to fixate the target (the male face in the MTG and the face next to the singleton color dot in the DTG). The stimuli remained on screen until the program recorded a valid target fixation. If participants did not fixate the target within 2000 ms after display presentation, the trial was ended. In this case, a feedback message “Target nicht gefunden” (target not found) was presented on screen for 1000 ms. Feedback was included to motivate participants to search for the target and to avoid misunderstandings of the instructions. After an intertrial interval of 500 ms, the next trial started. Participants started with 20 practice trials (with the face orientation with which they started first) and then completed four blocks of 60 trials each.

7.2.5 Apparatus

The apparatus was the same as in Study 1 and in Study 3.

7.2.6 Eye Movement Event Detection

The same event detection criteria as in Study 1 and Study 3 were applied. A saccade or fixation was considered to fall on the target if it landed within an area of interest defined as a rectangle of 170 pixels width and 220 pixels height with its center being the stimulus center.

7.2.7 Overview of Outcome Variables

In contrast to the previous studies, participants did not give a manual response. Thus, only overt attentional variables were analyzed. The overt "counterpart" of manual response latency was s*earch time:* the time from the onset of the stimulus display until a saccade fell on the target. Correspondingly, *accuracy* refers to the percentage of trials in which the target was fixated in time. Except for this change, the same eye movement variables as in Study 1 and in Study 3 were analyzed.

7.2.8 Data Preparation

The datasets of two participants had to be excluded from the analysis: one person reported shortsightedness, but only after completion of the experiment, and another individual had a very high error rate (63%), which was classified as an outlier value in boxplot analysis. Thus, 21 persons were in the DTG and 18 persons in the MTG. To check whether the target detection procedure during the experiment was operating reliably, it was tested if a target saccade was missed by the procedure and if trials were falsely discarded as erroneous. For each participant, the procedure was working correctly in every trial: if the on-line target fixation check did not register a saccade and the trial was marked as an error, the recorded gaze did not fall on the target any time during the trial. 28% of the trials had to be discarded due to errors (i.e., the participant did not make a saccade on the target). The same trimming criteria as in the previous studies were applied: Only 0.05% of the trials had to be discarded because of too slow or too fast search times. 6.1% of the trials had to be excluded from the dataset because the first saccade was an anticipatory one (0 ms > saccadic latency < 80 ms), another 0.1% because the eye tracker did not record any saccade during the trial (saccadic latency = 0 ms) and 0.3% due to a very slow saccadic latency (z-transformed saccadic latency > 3.29). Overall, 35% of the trials had to be excluded from the dataset due to these criteria.

7.3 Results

Based on the hypotheses the following a-priori contrasts for the independent variable trial type were defined:

- Contrast 1 (C1) compared the all neutral trial type to the two emotional target trial types (happy target and angry target).
- Contrast 2 (C2) compared the all neutral trial type to the two emotional distractor trial types (happy distractor and angry distractor).
- Contrast 3 (C3) compared the angry target trial type to the happy target trial type.
- Contrast 4 (C4) compared the angry distractor trial type to the happy distractor trial type.

Contrasts 1 and 2 are not orthogonal. Nevertheless, they were employed because the comparison to the neutral baseline is necessary to test whether a distraction and/or facilitation effect is present in the data. In case of significant results for the two contrasts, additional Bonferoni-Correction is applied to control for α–error-inflation due to non-orthogonal contrasts.

7.3.1 Number of Fixations

The number of fixations in the DTG was significantly lower than 2, $t(20) = -4.84$, $p < .001$, 95% CI [-0.64, -0.26], and participants needed on average 1.55 fixations to find the target (95% CI [1.36, 1.74], range [0.69, 2.18]). In the MTG, however, roughly 2.61 fixations were needed per trial (95% CI [2.39, 2.82], range [1.53, 3.15]). The number of fixations in the MTG was significantly higher than 2.0, $t(17) = 5.91$, $p < .001$, 95% CI [0.39, 0.82]. These results show that search was parallel in the DTG and serial in the MTG.

7.3.2 Accuracy

To test the hypotheses, ANOVAs with task goal (MTG, DTG) as a between-subjects variable and trial type (all neutral, happy target, angry target, happy distractor, angry distractor) and face orientation (upright, inverted) as within-subjects variables were conducted on all of the dependent variables. The ANOVA on accuracy revealed only a significant interaction between face orientation and task goal, $F(1, 37) = 12.38$, $p = .001$, $\eta^2_p = .25$. The remaining effects were not significant, *F*s < 2.71, *p*s > .10. In the DTG, accuracy did not differ significantly between the two face orientation conditions (*M* = 68% with upright faces, *M* = 72% with inverted faces), $F(1, 20) = 2.70$, $p = .12$, whereas in the MTG, participants' accuracy was significantly lower with inverted faces (*M* = 71%) than with upright faces (*M* = 78%), $F(1, 17) = 12.63$, $p = .002$, $\eta^2_p = .43$.

7.3.3 Search Time

The 2 (task goal) x 5 (trial type) x 2 (face orientation) – ANOVA on search time resulted in a large main effect of task goal, $F(1, 37) = 105.11$, $p < .001$, $\eta^2_p = .74$, with longer search times in the MTG and a main effect of face orientation, $F(1, 37) = 25.59$, $p < .001$, $\eta^2_p = .41$, with slower search times in the inverted faces condition. Both effects interacted, $F(1, 37) = 17.91$, $p < .001$, $\eta^2_p = .33$. As Figure 22 shows, the difference between the inverted and the upright faces condition was larger in the MTG than in the DTG. Furthermore, the main effect of trial type was significant for C2 (all neutral vs. emotional distractors), $F(1, 37) = 6.57$, $p = .015$, $\eta^2_p = .15$, and interacted with face orientation, $F(1, 37) = 5.40$, $p = .026$, $\eta^2_p = .13$. Finally, the three-way Face Orientation x Trial Type x Task Goal interaction was significant for C2, too, $F(1, 37) = 5.11$, $p = .03$, $\eta^2_p = .12$. The remaining effects did not reach significance, *F*s < 1.95, *p*s > .17.

Additional *t*-tests that compared the all neutral condition with the two distractor conditions for inverted and upright faces were computed in order to disentangle the significant Face Orientation x Trial Type x Task Goal interaction for C2. If an emotional distractor effect (i.e., a significant delay in search time in the emotional distractor condition compared to the all neutral condition) was solely due to stimulus salience, the distraction effect should be present in the inverted and in the upright faces condition. If, however, an emotional distractor effect was due to the emotional expression of the face, the distraction effect should be present only in the upright condition. In the DTG, these tests yielded non-significant results in the upright faces condition, *t*s < 1.55, *p*s > .13 and in the inverted faces condition, *t*s < 1.88, *p*s > .07. Thus, none of the emotional distractors delayed search time in the upright or in the inverted faces condition. In the MTG, however, search time was significantly delayed by 86 ms in the happy distractor condition, $t(17) = 2.54$, $p = .021$, 95% CI [15, 158] and by 115 ms in the angry distractor condition, $t(17) = 3.25$, $p = .005$, 95% CI [40, 189] with upright faces. Inverted emotional faces did not cause any distraction, *t*s < 1 (see Figure 22).

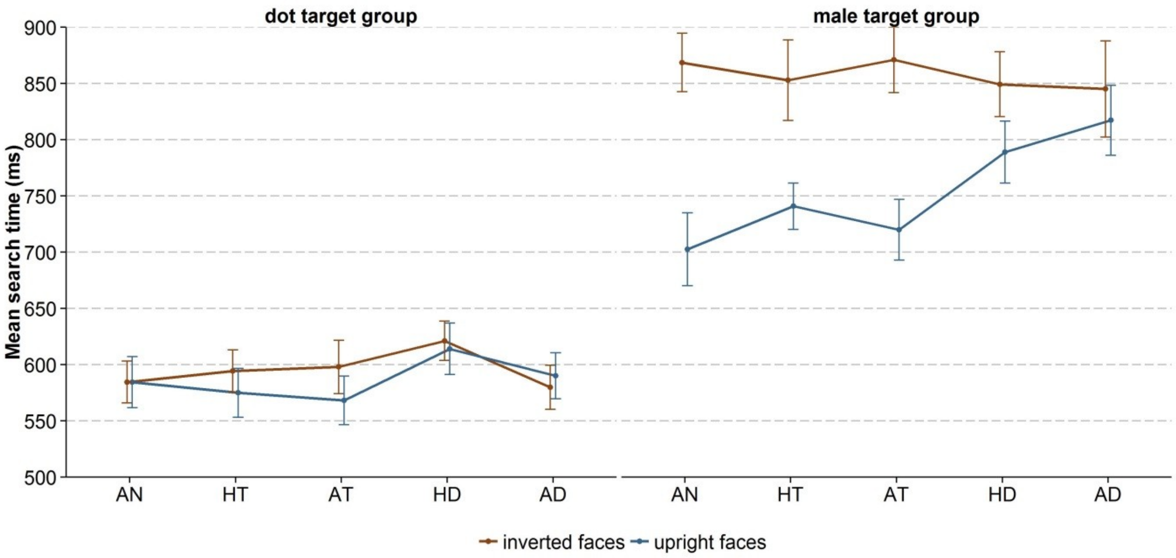

Figure 22. Mean search times, separately for the two task goal groups, dependent on trial type (AN = all neutral; HT = happy target; AT = angry target; HD = happy distractor; AD = angry distractor) and face orientation. Error bars represent standard errors. Points are set off horizontally to enhance the visibility of the error bars (Study 4).

7.3.4 Saccadic Latency

As the first saccade landed on the target only in 14 - 20% of the trials (see Chapter 7.3.5), only 17 participants had data in each of the ten cells of the within-subjects variables. Therefore, saccadic latency was not analyzed as a dependent variable in this analysis.

7.3.5 Location of First and Later Saccades

Like in the other studies, the random probability to hit one of the stimuli with the first saccade was approximately 17% (if one selects one of the six stimulus positions as a definite target). The hit rates of the first and of all saccades on the different stimuli (target, neutral distractor face and emotional distractor face) are displayed in Table 10. Concerning the first saccades, Table 10 shows that emotional distractor faces were not hit more often than neutral distractors and that the hit rate lies below the random hit rate.

The hit rate of the first saccade on the target was, at least in some trial types, above the random probability, but only in the DTG. Concerning the hit rate of all saccades, the emotional distractor was fixated sometime during the trial in 20 - 24% of the trials in the DTG, and in 40 - 41% of the trials in the MTG. These values are markedly higher than those for the neutral distractors (M = 14 - 19%) and thus provide a first hint at what the cause for the delay in search time is. It probably is not due to an early capture of oculomotor attention, but due to interference later in attentional shifting. In the next two chapters (7.3.6 and 7.3.7), two ANOVAs tested whether emotional distractors diverted attention away from the target and thus decreased the hit rate of the first saccade on the target and whether emotional faces captured attention to a greater extent than neutral distractors.

Table 10. *The Hit Rates of the First Saccade and of All Saccades on Different Stimuli (SD in Brackets), Separately for the two Task Goal Groups, Dependent on Trial Type and Face Orientation (Study 4).*

		DTG						MTG					
		First saccade			All saccades			First saccade			All saccades		
		TA	ND	ED	TA	ND	ED	TA	ND	ED	TA	ND	ED
AN	UP	22	10	---	68	14	---	16	10	---	77	15	---
		(13)	(3)		(19)	(3)		(9)	(4)		(13)	(3)	
	IN	19	10	---	73	14	---	13	12	---	71	16	---
		(12)	(4)		(17)	(3)		(9)	(3)		(15)	(2)	
HT	UP	26	10	---	69	13	---	11	12	---	79	17	---
		(15)	(3)		(16)	(3)		(7)	(3)		(14)	(1)	
	IN	19	10	---	72	13	---	16	10	---	73	16	---
		(15)	(3)		(17)	(3)		(10)	(4)		(15)	(2)	
AT	UP	20	9	---	68	13	---	13	10	---	75	15	---
		(11)	(3)		(18)	(3)		(7)	(3)		(14)	(2)	
	IN	18	11	---	72	14	---	10	12	---	72	17	---
		(15)	(3)		(16)	(3)		(10)	(3)		(16)	(2)	
HD	UP	17	10	13	65	15	23	15	11	15	79	19	41
		(11)	(3)	(12)	(17)	(3)	(16)	(10)	(4)	(9)	(17)	(3)	(14)
	IN	19	12	7	72	15	20	14	11	16	72	18	41
		(13)	(5)	(8)	(16)	(4)	(11)	(12)	(3)	(11)	(14)	(4)	(16)
AD	UP	21	10	10	69	16	21	13	11	9	77	19	42
		(14)	(4)	(7)	(19)	(4)	(13)	(8)	(5)	(8)	(16)	(3)	(16)
	IN	19	10	12	69	14	24	17	11	13	66	17	41
		(15)	(4)	(11)	(17)	(4)	(10)	(14)	(4)	(8)	(13)	(4)	(16)

Note. AN = all neutral; HT = happy target; AT = angry target; HD = happy distractor; AD = angry distractor; UP = upright face; IN = inverted face.

7.3.6 Hit Rate of the First Saccade on the Target

The 2 (task goal) x 5 (trial type) x 2 (face orientation) – ANOVA on the hit rate of the first saccade on the target yielded a main effect of trial type for C3 (happy target vs. angry target),

$F(1, 37) = 5.63$, $p = .023$, $\eta^2_p = .13$, which was modulated by an interaction with face orientation and task goal, $F(1, 37) = 5.79$, $p = .021$, $\eta^2_p = .14$. In addition, there was a significant main effect of task goal: Participants in the DTG made a correct target saccade more often (M = 20%) than participants in the MTG (M = 14%), $F(1, 37) = 6.75$, $p = .013$, $\eta^2_p = .15$. The remaining effects were non-significant, Fs < 2.03, ps > .16.

To disentangle the three-way interaction between trial type, face orientation and task goal for C3, separate t-tests comparing the angry target and the happy target condition were computed in the two task goal groups. In the DTG, the hit rate on the happy target was higher than that on the angry target in the upright condition, $t(20) = 2.30$, $p = .032$, 95% CI [.01, .12], but not in the inverted condition, $t < 1$. In the MTG, the happy target was the target of the first saccade more often than the angry target in the inverted condition, $t(17) = 2.28$, p = .036, 95% CI [.01, .13], but not in the upright condition, $t < 1$. Yet, in both groups and in both face conditions, the hit rates did not differ between the angry or happy target and the all neutral baseline, ts < 1.58, ps > .13 (see Figure 23).

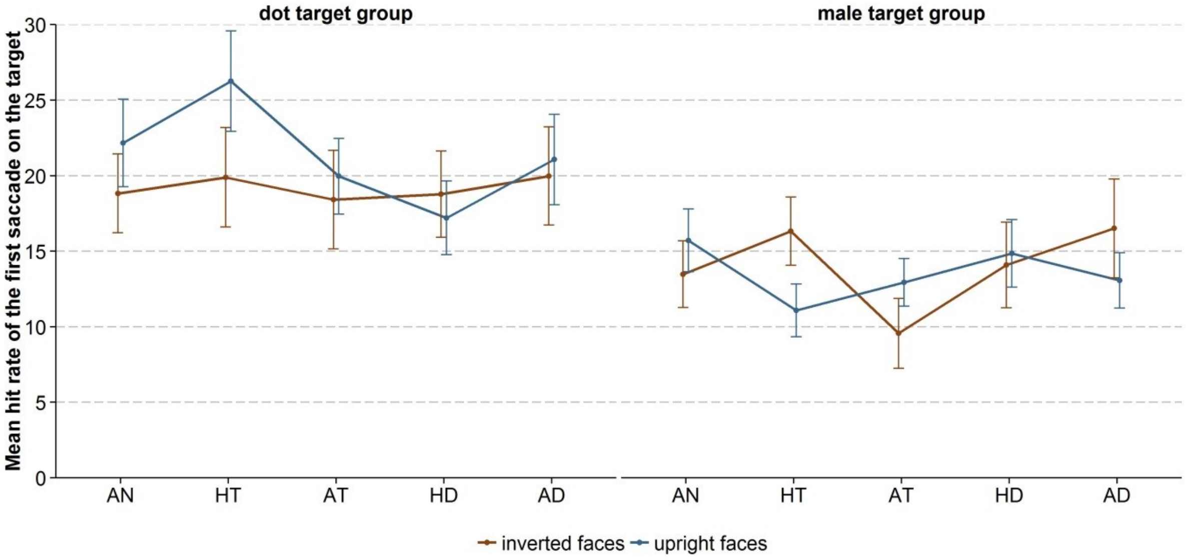

Figure 23. Mean hit rate of the first saccade, separately for the two task goal groups, dependent on trial type and face orientation. Error bars represent standard errors. Points are set off horizontally to enhance the visibility of the error bars (Study 4).

7.3.7 Comparison of the Hit Rate of the First Saccade on the Emotional Distractor to the Hit Rate on the Neutral Distractors

An ANOVA compared the hit rate of the first saccade on the emotional distractors to that on neutral distractors. This analysis included saccade target (emotional distractor, neutral distractor), face orientation (upright, inverted), trial type (happy distractor, angry distractor) and task goal (DTG, MTG) as independent variables. It revealed only a two-way Face Orientation x Trial Type interaction, $F(1, 37) = 4.63$, $p = .038$, $\eta^2_p = .11$ (in the upright condition, both emotional expressions were equally often the target of the first saccade, but in the inverted condition, happy distractors were fixated more often than angry distractors) and a three-way Trial Type x Saccade Target x Task Goal interaction, $F(1, 37) = 5.13$, $p = .029$, $\eta^2_p = .12$. The remaining effects were non-significant, Fs < 3.51, ps > .06.

To disentangle the significant three-way interaction Trial Type x Saccade Target x Task Goal, separate ANOVAs for each task goal group were conducted. In the MTG, the interaction between trial type and saccade target was significant, $F(1, 17) = 5.68$, $p = .029$, $\eta^2_p = .25$. This effect indicated that the hit rate on happy emotional distractors was significantly higher than the one on neutral distractors, but that the hit rates on angry distractors and on neutral distractors did not differ. As this effect was not modulated by face orientation, it probably reflects bottom-up attentional capture of the white teeth of the happy distractor. In the DTG, the Trial Type x Saccade Target interaction was not significant, $F < 1$.

7.3.8 Fixation Duration on Emotional Distractors

As the previous two chapters did not reveal a reason for the distractor effect by emotional upright faces in search time, the delayed disengagement hypothesis (Belopolsky et al., 2011; Fox et al., 2001; Fox et al., 2002) was tested. According to this hypothesis, an emotional distractor delays the disengagement of attention once it is fixated. Therefore, the fixation duration on the emotional distractor should be longer than the one on a neutral distractor. An ANOVA with trial type (happy distractor, angry distractor), face orientation (upright, inverted), fixation target (emotional distractor, neutral distractor) and task goal (DTG, MTG) yielded significant main effects of trial type, $F(1, 37) = 12.13$, $p = .001$, $\eta^2_p = .25$, of task goal, $F(1, 37) = 8.59$, $p = .006$, $\eta^2_p = .19$, and of face orientation, $F(1, 37) = 4.21$, $p = .047$, $\eta^2_p = .10$, with longer fixations on angry distractors ($M = 148$ ms vs. $M = 137$ ms), in the MTG ($M = 153$ ms vs. $M = 132$ ms) and on inverted faces ($M = 148$ ms vs. $M = 137$ ms). The main effect of fixation target was clearly not significant, $F < 1$. The remaining effects were non-significant, Fs < 2.03, ps > .16.

7.3.9 Number of Fixations – Further Analysis

As the previous analyses did not reveal a possible cause for the delay in search time, the number of fixations was analyzed as a dependent variable, too. The 2 (task goal) x 5 (trial type) x 2 (face orientation) - ANOVA on the number of fixations during a trial yielded a significant main effect of face orientation, $F(1, 37) = 8.32$, $p = .007$, $\eta^2_p = .18$ and a main effect of task goal, $F(1, 37) = 58.50$, $p < .001$, $\eta^2_p = .61$, with more fixations in the inverted condition and in the MTG. Furthermore, there was a significant two-way interaction between face orientation and trial type for C2, *F*(1, 37) = 4.89, *p* = .033, η^2_p = .12. Importantly, like for search time, this two-way interaction was further qualified by an interaction with task group, *F*(1, 37) = 4.77, *p* = .035, η^2_p = .11. The data pattern in Figure 24 looks very similar to the pattern in Figure 22. The remaining effects were not significant, *F*s < 2.30, *p*s > .13.

T-Tests revealed that the angry and the happy distractor did not lead to significantly more fixations in the DTG, neither in the upright nor in the inverted condition, *t*s < 1.39, *p*s > .17. In the MTG, however, participants made more fixations in the angry distractor trials (*M* = 2.68) than in the all neutral trials (*M* = 2.37) with upright faces, *t*(17) = 2.26, *p* = .037, 95% CI [0.60, 0.02] and the difference between the happy distractor condition (*M* = 2.60) and the all neutral baseline (*M* = 2.37) for upright faces was marginally significant, *t*(17) = 1.88, *p* = .078, 95% CI [0.50, -0.03]. In the inverted faces condition, however, the number of fixations did not differ between the baseline and the distractor conditions, *t*s < 1.40, *p*s > .17.

To check if the increase in the number of fixations in these conditions was due to fixations on the emotional distractors, the number of fixations on the emotional distractors in the upright condition was compared to the number of fixations on the emotional distractors in the inverted condition in the MTG, with the assumption that the number of fixations would be higher in the former. Yet, this was not the case. The number of fixations on happy upright faces (*M* = 0.47) did not differ from that on happy inverted faces (*M* = 0.56), *t*(17) = -1.39, *p* = .18, 95% CI [-0.23, 0.05], and the number of fixations on angry upright faces (*M* = 0.50) did not differ from that on angry inverted faces (*M* = 0.59), *t*(17) = -1.25, *p* = .23, 95% CI [-0.27, 0.07]. Therefore, the increase in the number of fixations in the MTG for upright faces was not due to an increase in emotional fixations, compared to the upright faces condition. The data rather suggest that participants in the MTG "compensate" the fixation on the inverted emotional distractor with fewer fixations on other locations. As a result, the number of fixations does not differ between the neutral baseline and the emotional distractor conditions. In the upright faces condition, however, participants make an *additional* fixation to the emotional distractor, compared to the number of fixations in the neutral baseline. This was confirmed by the fact that in the inverted faces condition, participants made significantly fewer fixations on neutral distractors in the happy distractor condition (*M* = 1.86), *t*(17) = 4.42, *p* < .001, 95% CI [0.30, 0.86] and in the angry distractor condition (*M* = 1.67), *t*(17) = 9.18, *p* < .001, 95% CI [0.59, 0.95], compared to the neutral baseline (*M* = 2.44). In the upright condition, however, participants did not make fewer fixations on neutral distractors in the happy (*M* = 1.76), *t*(17) = 1.44, *p* = .17, 95% CI [-0.08, 0.40]), and in the angry distractor condition (*M* = 1.76), *t*(17) = 1.33, *p* = .20, 95% CI [-0.10, 0.42], compared to the baseline (*M* = 1.92). The same was true for the number of fixations that landed in between the stimuli for the upright condition. This neither differed between the baseline (M =0 .44) and the happy distractor condition (M = 0.37), $t(17)$ = 1.95, p = .068, 95% CI [-0.01, 0.14], nor between the baseline and the angry distractor condition (M = 0.43), t < 1.

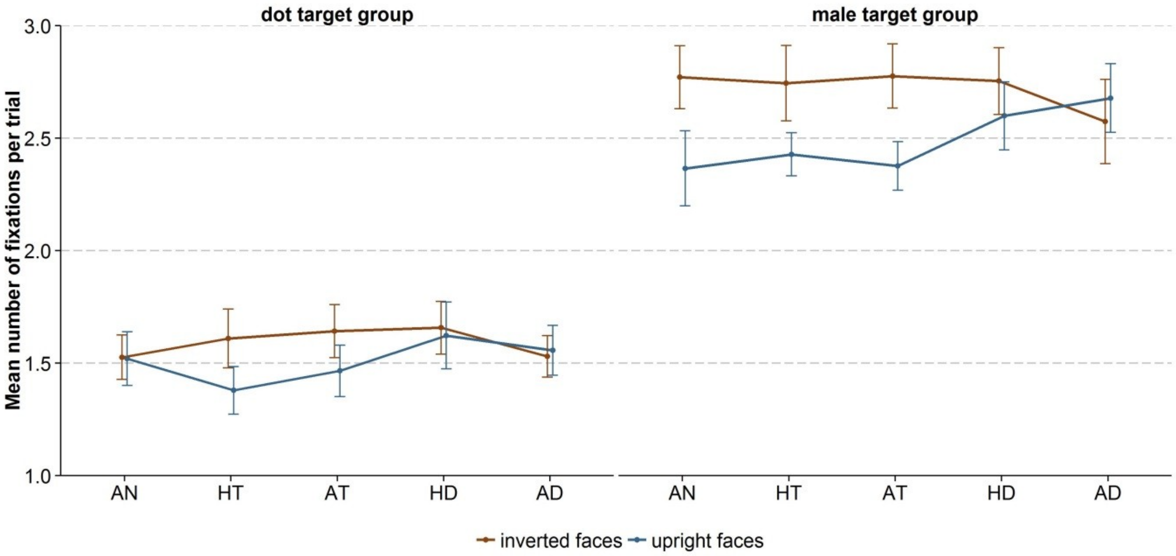

Figure 24. Mean number of fixations per trial, separately for the two sequence groups, dependent on trial type and face orientation. Error bars represent standard errors. Points are set off horizontally to enhance the visibility of the error bars (Study 4).

7.3.10 Gender Effects

In the MTG, the target was always male and all of the distractors were female, whereas target and distractor gender were varied in the DTG. Therefore, only trials with a male target and female distractors (resembling the target-distractor gender combination in the MTG) were analyzed separately for participants in the DTG. The datasets of five participants had to be discarded from this analysis because the mean search time in one condition could not be computed due to a lack of correct trials in this condition. The 5 (trial type) x 2 (face orientation) - ANOVA was conducted with the 16 remaining datasets and revealed non-significant effects for all factors and contrasts, Fs < 2.39, ps > .14.

7.3.11 The Cause of the Emotional Distractor Effect in Search Time

The results showed that emotional upright faces delay search time only in the MTG. The analyses of the hit rates of the first saccade revealed that these were not affected specifically by upright emotional distractors in this group. Furthermore, the eyes did not fixate longer on upright emotional distractors than on neutral distractors. Therefore, a delay of attentional disengagement could not be the cause. Instead, the delay in search time occurred because participants made on average 0.23 (in the happy distractor condition) / 0.31 (in the angry distractor condition) additional fixations – compared to the neutral baseline – in emotional distractor trials. This difference in the number of fixations was only significant between the angry distractor condition and the neutral baseline, yet the difference between the happy distractor condition and the neutral baseline was close to significance. One could argue that a happy distractor led to less fixations on "other" locations (this difference was also marginally significant), whereby a potential additional fixation on emotional distractors was compensated. Yet, a look at the numbers reveals that this would only amount to 0.07 fewer fixations in the happy distractor condition than in the baseline, whereas the difference between the total numbers of fixations between these conditions is 0.23. Thus, the most obvious explanation for the delay in search time is that participants made, in some trials, an additional fixation on the emotional distractor while searching for an upright target and therefore needed more time to make a saccade to the target. As the mean fixation duration on an upright emotional distractor in the MTG lasts around 138 ms (happy distractor) / 155 ms (angry distractor), participants' fixation duration is about 32 ms (happy distractor) / 48 ms (angry distractor) longer in the distractor trials. Each additional fixation is of course accompanied by two additional saccades (one to the distractor, another from the distractor to the target), each of which takes at least an additional 30 ms. These processes explain why participants need about 86 ms (happy distractor) or 115 ms (angry distractor) longer to saccade to the upright target when an upright happy or angry distractor is present.

7.4 Discussion

The present study tested three hypotheses. It was assumed that emotion does not capture attention when the task set allows for parallel search (Hypothesis 4.1), but interferes with visual search when faces need to be fixated in order to detect their gender (Hypothesis 4.2). Consequently, it was expected that participants in the DTG do not show any signs of emotional attentional capture, whereas emotional distractors should interfere with search in the MTG. This interference was expected to be larger for happy faces than for angry faces (Hypothesis 4.3). The design also allowed testing facilitation effects of emotional faces.

Based on Hypothesis 4.1 and Hypothesis 4.2, it was expected that search should be parallel in the DTG and serial in the MTG. This assumption was confirmed by the results. Participants in the DTG needed on average 1.6 fixations to find the target, which was significantly lower than two. This is consistent with the result of Study 1, in which the number of fixations was also significantly lower than two. Furthermore, Study 2 reported flat search slopes for the same search task and with the same material as it was used in Study 1. This verified the conclusion that search with less than two fixations is parallel. In the MTG, however, roughly 2.6

fixations were needed per trial and the number of fixations in the MTG was significantly higher than two. This is similar to the number of fixations in Study 3 and suggests that search was serial when the male face among female faces had to be found. Furthermore, the mean fixation duration on a face was significantly longer in the MTG (153 ms) than in the DTG (132 ms). This shows that participants in the MTG processed the faces deeper than participants in the DTG. Preattentive processes during search in the MTG could still be analyzed with the hit rate of the first saccade (like in Study 3).

The next three chapters (7.4.1 - 7.4.3) discuss whether the three hypotheses were confirmed by the results. Chapter 7.4.4 addresses some limitations of the present study and Chapter 7.4.5 draws a conclusion.

7.4.1 An Emotional Face Interferes with Visual Search…

Importantly, any interference effect in search time in the DTG would be a capture effect because it occurred during parallel search. Yet, as expected, this was not the case. In the DTG, inverted and upright emotional faces did not delay search time in comparison to the neutral baseline. A facilitation effect of a happy or an angry target was not observed, either. The percentage of first saccades landing on the target was significantly higher when the target was happy than when it was angry, yet it was not significantly higher than the percentage of first saccades landing on a neutral target, which is the relevant baseline. In the same vein, emotional distractors did not capture attention by being the target of the first saccade more often than neutral distractors.

In the MTG, however, participants needed more time to make a saccade to the target if an angry or a happy distracting face was present in the display. As this distraction effect was only observed with upright faces, it could be completely attributed to the emotional content of the faces. Further analyses confirmed Hypothesis 4.2, which assumed that an interference effect in the MTG is not observed in early attentional orienting, but either in a failure to disengage attention once an emotional face is fixated or due to a distraction in attentional shifts. The data for the hit rate of the first saccade on the emotional distractors showed that the probability to hit the emotional distractor was not higher in the MTG than in the DTG, and in both groups, emotional upright distractors did not draw more initial saccades than a neutral distractor. The proportion of first saccades on the target also was not significantly decreased when an upright emotional face was present. Therefore, the delay in search time by upright emotional faces in the MTG could not be attributed to preattentive capture. The probability to hit the emotional distractor with any saccade (the first, the second, or the third) was much higher in the MTG than in the DTG, but for both inverted and upright distractor faces. Fixation durations on upright emotional distractors were not significantly longer than fixation durations on upright neutral distractors, either. Consequently, a delay of attentional disengagement by the emotional distractor could not be the reason for the interference effect.

Thus, the overall number of fixations and the number of fixations on the different locations was analyzed. This resulted in the finding that participants in the MTG made more fixations in the upright emotional distractor conditions compared to the neutral baseline, whereas the number of fixations did not differ between these conditions for inverted faces. This pattern of results therefore reflects the results from search time and is the most likely explanation for the delay in search time by upright angry and happy faces. The fact that the number of fixations on emotional distractors did not differ between the upright and the inverted condition does not contradict this explanation. In the inverted condition, participants made fewer fixations on neutral faces when an emotional distractor was present. This demonstrates that an emotional inverted face was equivalent to a neutral face in this condition. When searching among inverted faces, participants needed approximately 2 – 3 fixations in order to find the target, irrespective of the emotional expressions on these faces. The important fact is that they did not need more fixations to find the target when an emotional distractor was present. When searching among upright faces, however, the emotional distractor had a different quality because it forced participants to make an *additional* fixation in some trials and did not lead to fewer fixations on neutral distractor faces.

To conclude, these results demonstrate that participants' gaze was distracted by the emotional face: in some trials, the eyes additionally fixated on it, but if it was fixated, it was not processed for a longer time than a neutral face. The evidence that this effect was due to the emotional content of the stimuli is strong because the delay was not observed at all with upright faces. Thus, this evidence is more convincing than the results by Huang et al. (2011), which only showed a decrease in interference (the authors called it capture) with inverted faces, which means that the effect was driven, to a considerable extent, by low-level confounds of the emotional faces. This might also indicate that the inversion of schematic faces has a different effect on face processing than the inversion of photographic faces.

7.4.2 …But does not Capture Attention

The results of Hodsoll et al. (2011) and, to a lesser extent, of Huang et al. (2011), suggested that emotional faces, and angry faces in particular, capture attention preattentively. It was argued that both studies did not provide unequivocal evidence for emotional capture because they did not provide data on the efficiency of search (whether search was parallel or serial) and they did not rule out all possible low-level confounds. In contrast to these studies, the present study employed an inverted condition to check for visual confounds, measured eye movements to detect the cause for any interference or capture and compared two task goals to test whether emotion captured attention (preattentively) or interfered with attentional shifting (during the serial stage). Although set size was not varied (which was not varied in the other studies, either), the recording of eye movements allowed making an inference about the underlying search process. The fact that participants needed 1 – 2 saccades to find the target strongly suggests that search was parallel in the DTG. As the emotional distractor face did not affect any indicator of attentional capture in the DTG, Hypothesis 4.1 was clearly confirmed. Another aspect that underlines the validity of this conclusion is the fact that, in contrast to many other studies on this issue (one notable exception is the study by Calvo et al. (2008)), 38 different facial identities were used as stimuli. Thus, participants could not detect the target by learning subtle feature differences between target and distractor faces (cf. Becker et al., 2011). Consequently, the results of Study 4 demonstrate that emotions do not automatically capture attention, but only interfere with search if emotion processing is associated with the task set – in this case identifying the face gender.

7.4.3 Emotion-Superiority Instead of Threat-Superiority or Happiness-Superiority

Based on the results by Becker et al. (2011), it was expected that happy faces distract attention to a greater extent than angry faces (Hypothesis 4.3). This hypothesis was clearly disconfirmed by the results because search time was delayed both in the angry and in the happy distractor condition in the MTG, without a significant difference between the two. The same pattern of results was found for the number of fixations. There was only a tendency of a happiness-superiority effect in the DTG because participants made a greater proportion of first saccades towards the happy target than towards the angry target. However, the hit rate on the happy target did not differ from the hit rate on the neutral target, so that this result does not reflect a superior detection of happy targets, but most probably rather a random finding. The result that both angry and happy faces distract attentional processing is in line with the results of Hodsoll et al. (2011), the important difference being that Hodsoll et al. argue that emotional faces capture attention preattentively. Furthermore, other studies using the odd-one-out task instruction ("Find the discrepant face.") have employed up to six different emotional faces and found detection advantages for several emotions (Calvo et al., 2008; M. A. Williams et al., 2005). The present study only compared two emotions. Further research studying all six basic emotional expressions (surprise, anger, fear, sadness, disgust, happiness) should be carried out with this paradigm to check whether any irrelevant emotion distracts attentional processing during serial search or if this effect is restricted to angry and happy faces.

As a conclusion, the results of the present study challenge the fear module account (Öhman & Mineka, 2001) and the threat superiority hypothesis (Öhman et al., 2001), which argue that threatening emotions capture attention preattentively: Firstly, none of the emotional faces distracted attention during parallel search, and secondly, both threatening and happy faces slowed search time.

7.4.4 Limitations

Although the results provide convincing evidence for emotional interference of visual search, some conclusions might be put into question by limitations of the study. One possible confound was already addressed in Chapter 7.3.10. As the target was always a male face in the MTG, and as target gender was varied in the DTG, one could argue that the interference effect in the MTG is driven by female emotional expressions. Therefore, only trials with a male target and female emotional distractors were analyzed separately in the DTG. This revealed only non-significant results. The validity of this analysis is a bit qualified by the fact that the data of five participants could not be included in the analysis because they lacked data in at least one cell of the conditions. Therefore, further research is needed in order to investigate gender effects in parallel and serial search tasks. It would be highly interesting to conceptualize different task goals that call for either parallel or serial search while keeping all of the other attributes in the search display constant. This could be implemented, for example, by cueing participants with the target of search in each trial. Such a procedure would allow for a flexible combination of different identities and genders in the display. On the other hand, it has to be taken into account that variable target search is in general more difficult (see Study 3).

Another limitation pertains to the conclusion that search was parallel in the DTG and serial in the MTG. The claim for parallel search in the DTG is supported by the results of Study 1 and Study 2, which showed that 1 - 2 fixations per trial are typical for parallel search in an additional singleton task. This is consistent with previous studies showing that, if more than two saccades are needed, search is serial (Findlay & Gilchrist, 2003; Scialfa & Joffe, 1998; D. E. Williams et al., 1997; Zelinsky & Sheinberg, 1997). The conclusion that search for the male face among female faces operated with a serial sequence is supported by the theories of visual search that argue that neither faces nor emotional expressions conveyed by faces can be preattentively processed (cf. Wolfe & Horowitz, 2004). As the manipulation of set size is still the most widely used method to make inferences about the search process, it is necessary to further confirm the conclusion that search was serial in the MTG by varying set size and monitoring eye movements in another study. This seems especially important because participants' error rate was higher in Study 1, Study 3 and Study 4 than in the studies cited that investigated differences in parallel and serial search with response latencies and eye movements. The mean error rate was less than 3% for parallel search tasks and up to 15% for conjunction search tasks in Scialfa and Joffe (1998), Zelinsky and Sheinberg (1997) and D. E. Williams et al. (1997). More research is needed to examine whether the cut-off of two fixations is also applicable to search tasks with a mean accuracy between 70% and 90%. Furthermore it is important to note than none of the studies that examined eye movements in parallel and serial search employed the additional singleton task. More research on this topic is needed, too. However, eye movement recording allowed directly measuring the preattentive components of search, irrespective of whether the whole process of search was serial or parallel. Therefore, it was possible to examine whether emotion affected parallel processing in the MTG and in the DTG. As the results for the hit rate of the first saccade showed, this clearly was not the case in both groups. The issue of differentiation between parallel and serial search is further discussed in Chapter 8.

A third limitation concerns the error rate. The error rate of this study is higher (30% in the DTG and 26 % in the MTG) than the one of other studies on emotional capture that measured manual response time (e.g., Hodsoll et al., 2011). However, this is probably an inadequate comparison. In the latter study, participants had to indicate the orientation of the target face (right or left), which meant that the random probability to give the correct answer was 50%. In the present study, the random probability to give the correct answer was 17% because the task was to fixate the target. In addition, the study by Devue et al. (2012), whose task instructions most closely resemble the one in the DTG, does not report accuracy in text, but a line graph indicates that the mean accuracy for the upright faces condition amounts to approximately 60%. More research with these "fixate-the-target" tasks is needed to determine some kind of "standard" accuracy in this paradigm.

7.4.5 Conclusion

To conclude, the results of Study 4 confirmed Hypothesis 4.1 and Hypothesis 4.2: Emotional faces do not capture attention automatically, but distract attentional orienting if the detection of emotional expressions is associated with the top-down task goal. This conclusion challenges accounts that argue that emotional expressions capture attention without attention having to focus on them (Hodsoll et al., 2011; Huang et al., 2011; Öhman et al., 2001) and supports the cognitive theories of visual search that claim that only basic features can be preattentively processed (cf. Wolfe & Horowitz, 2004). Further research is needed to corroborate the results of Study 4 and to test whether the effects can be replicated with different tasks and stimuli. The result concerning Hypothesis 4.3 was surprising because the results of Becker et al. (2011) delivered strong support for a happiness-superiority-effect. Differences in the task set might explain this discrepancy: Whereas participants in the studies of Becker et al. (2011) were instructed to search either for the discrepant face or for a face with a certain emotional expression, observers in Study 4 (MTG) had to search for the male face. Consequently, the effect of certain emotional expressions on attentional orienting might be dependent on the specific top-down task set. The fact that both angry and happy faces drew additional fixations in the MTG in Study 4 suggests that any emotional expression is in some way relevant when the task goal is to identify the gender of individuals. Therefore, it is very important to test whether this emotion-superiority effect occurs in a serial search task with a different task goal, too.

Previous research has neglected the role of task goals in search tasks with emotional faces: Studies that systematically manipulate various aspects of the task set are needed to foster research on the topic of emotional distraction during visual search.

8 DISCUSSION

This chapter first summarizes the key findings of the four experimental studies and outlines their theoretical implications (see Chapter 8.1) and their limitations (see Chapter 8.2). Based on the results of the studies, some ideas and challenges for future research are discussed (see Chapter 8.3) and a final conclusion is drawn (see Chapter 8.4).

8.1 Summary of Key Findings and Theoretical Implications

This doctoral thesis examined the interplay of motivation and top-down control in visual search with the additional singleton task. Studies 1 to 3 addressed the hypothesis that attentional capture in the additional singleton task (see Chapter 2.1.3) is not a cognitive deficit of top-down control but the result of a motivational deficit - participants are not very motivated to suppress the salient distractor in the usual setting. It was assumed that participants regard the goal of distractor suppression more important and therefore invest more mental effort into top-down control if they receive an incentive for the fast and correct detection of the target. This assumption was based on the dimension weighting account (Müller et al., 2009; Zehetleitner et al., 2012; see Chapter 2.1.5), which argues that attentional capture is top-down modulable if participants invest sufficient effort, on cognitive energetics theory (Kruglanski et al., 2012; see Chapter 2.2.1), which proposes that goal importance is a determining factor of motivation and on experimental studies showing that incentives enhance cognitive control (e.g., Braver et al., 2014; see Chapter 2.2.2). The design and the stimuli of Study 1 were similar to the ones used in a previous experiment (Müller et al., 2009). Trials in which participants could win or lose, based on their performance, were compared to trials in which rewards or losses were not possible. The key finding of Study 1 was that saccadic latency and response time were decreased in incentive conditions compared to the neutral baseline and that participants did not show corresponding accuracy deficits, which means that they overall performed better if they had the chance to win or the possibility to avoid a loss. In the same vein, phasic pupil dilation was larger in incentive trials than in neutral trials. As the number of fixations was significantly below two, search was parallel in this study. Thus, ***incentives increased the transient mental effort invested in the task and enhanced top-down control during parallel search.***

Study 1 is - to the author’s knowledge - the first study that has demonstrated that performance-contingent rewards facilitate overt and covert top-down control in parallel visual search. So far, the study by Kiss et al. (2009) is the only other study in which performance-contingent rewards were found to enhance target detection in parallel visual search. In contrast to Study 1, the former employed a pop-out search task, in which the target was clearly the most salient item in the display. Therefore, the authors argued that reward neither affected top-down nor bottom-up control, but served as some kind of third control process. Study 1 employed a paradigm in which the CSD was the most salient item in the display, but reward facilitated search for the target nevertheless. Thus, it can be concluded that reward increased the top-down control during the task. This finding is in line with many previous studies showing that performance-contingent rewards enhance cognitive control in the stroop task and in the AX-task (Chiew & Braver, 2013; Fröber & Dreisbach, 2014; Locke & Braver, 2008; Savine et al., 2010; Savine & Braver, 2010; Veling & Aarts, 2010). Consequently, Study 1 delivered evidence that top-down control in visual search is indeed conceptually similar to cognitive control measured in response conflict tasks, as suggested by Müller et al. (2009). Furthermore, Study 1 showed that a payoff scheme that differed in some aspects from the payoff schemes of previous studies also had a marked influence on top-down control and mental effort. In comparison to many previous studies (see Table 1 in Chapter 2.2.2), participants were under more pressure in Study 1 because they were told that they would only receive the amount of money that they would win with their own performance and because their accumulated amount of money was communicated to them after each trial. It would be worthwhile to systematically manipulate different payoff schemes in an experiment in order to test whether their effect differs: One could compare a condition similar to the payoff scheme employed by Braver and colleagues (e.g., Chiew & Braver, 2013), in which participants receive a bonus payment (i.e., in addition to a fixed monetary fee) for good performance, to a condition similar to the payoff scheme employed in this thesis, in which participants were told that they only receive the amount of money that they win with their performance.

The most surprising result of Study 1 was that ***attention was not captured by the CSD in the neutral condition***. The most likely reason for this absence of attentional capture is that the effect of incentives spilled over to all of the conditions: Participants learned to suppress the CSD in the incentive conditions and transferred this practice to the neutral condition. Furthermore, results showed that attention was captured in the practice block at the beginning of the experiment. Therefore, it was assumed that participants learned to suppress the distractor effectively so fast that the design of Study 1 was not sensitive enough to track the changes in distractor interference. The unexpected results of Study 1 were compared to recent findings revealing that attentional capture is not a constant effect that is present during the whole experiment, but that participants are only distracted by the CSD in roughly half the trials of an experiment. Vatterott and Vecera (2012) reported that participants only need 24 trials to learn to suppress differently colored CSDs when they adopt feature search mode. In order to induce feature search mode, the authors used a display with heterogeneous distractors. Study 1, however, employed a display with homogeneous distractors that typically induces singleton search mode. Based on the dimension weighting account (Müller et al., 2009; Zehetleitner et al., 2012), which considers distractor practice as another determining factor (next to incentives) of top-down control over attentional capture, and the results of Study 1, it was assumed that the pattern of results uncovered by Vatterott and Vecera (2012) should also be evident with singleton search mode. This hypothesis was tested in Study 2.

Thus, Study 2 aimed at investigating practice effects occurring in the additional singleton task without manipulating performance-contingent incentives. Furthermore, Study 2 employed three different display sizes in order to confirm the finding from Study 1 that search was parallel. Like in the study by Vatterott and Vecera (2012), the experiment consisted of a practice block and four consecutive test blocks with four differently colored CSDs. Thus, the CSD changed its color in the beginning of each block. Analysis revealed that participants, who searched among the same display as in Study 1, only showed attentional capture during the first half of each block. However, in block 3 they showed no attentional capture at all and in block 4 they were distracted in both sub-blocks. These findings were addressed in the discussion of Study 2 (see Chapter 5.4.2). Overall, the results complemented those of Vatterott and Vecera (2012) because they proved that distractor practice in the additional singleton task is effective not only in feature search mode, but also in singleton search mode. On the one hand, distractor practice was tied to the specific color of the CSD because participants were distracted by a new distractor color. On the other hand, Study 2 also reported indications of a general linear reduction of attentional capture across blocks. Although this reduction was only marginally significant, it might hint at the possibility that distractor practice is not only specific, but can also be transferred to new features. Participants tended to suppress a new feature faster if they had already learned to suppress a different feature before.

Study 2 provided convincing evidence that attentional capture, when expressed as the mean difference in response latencies in trials with a CSD and those without, is relatively strong, but that it actually occurs in only half of the trials, and in the feature display group it was observed in only one quarter of the trials (in the first half of the first block). This finding is consistent with the dimension weighting account which assumes that participants can reduce their amount of attentional capture if they have the possibility to practice distractor suppression and if they have the incentives to do so (Müller et al., 2009). Evidence for the search mode account (Bacon & Egeth, 1994; Leber & Egeth, 2006; see Chapter 2.1.4), an opposing theory of top-down control in the additional singleton task, was weak. In Study 2, participants in the singleton display group were more distracted by the CSD than those in the feature display group, but they learned quickly to suppress the specific distractor and showed a tendency towards general distractor learning. Both processes are more consistent with the dimension weighting account. Thus, the results of Study 2 ***demonstrate that attentional capture is attenuated by distractor practice even if participants adopt singleton search mode.***

The results of Study 2 were analyzed with respect to the question of how a study should be designed that aims to examine the effect of incentives on top-down control in the additional singleton task. The first conclusion was that the singleton distractor color should change in each block, like in Study 2, so that stable distractor interference is observable at least in the first half of each block. The second conclusion concerned the implementation of incentives. Based on the fast reduction of attentional capture in Study 2 a block-wise manipulation was considered more adequate. The third conclusion, based on the tendency of a linear decrease of attentional capture over the blocks in Study 2, stated that a more difficult version of the additional singleton task with variable target and distractor forms should be used. This kind of task was expected to ensure a high level of attentional capture throughout the experiment. Attentional capture should be as large in the last block as in the first to allow aggregating mean values over the four blocks.

The results of Study 3 showed that search was different from Study 1 in one important aspect: Whereas search in Study 1 was parallel - as was confirmed by the results of Study 2 - search in Study 3 was serial, as indicated by a relatively high number of fixations executed during a trial. The more difficult search task obviously not only increased overall distractor interference, but also forced participants to engage in serial search. Thus, response latencies and accuracy could not be treated as indicators of attentional capture and only the proportion of first saccades towards the target and towards the CSD could be analyzed in this regard. Both were not modulated by incentive. The hit rate on the target was not higher in incentive trials compared to neutral trials and the hit rate on the CSD was not lower in incentive trials compared to neutral trials. In the same vein, global indicators of distractor interference (response latencies and accuracy) were not decreased by reward, either. Consequently, in contrast to Study 1, ***performance benefits in reward conditions were not observed***. ***Neither attentional capture nor distractor interference (as indicated by response latencies and accuracy) were decreased by incentives***. Yet, those individuals who started with a reward block, in comparison to those who started with a neutral block, learned faster to find the target efficiently. This effect was observed both with overt indicators of attentional capture and with response latency. It can be regarded as an indirect effect of incentive on distractor suppression: ***If participants' initial experience with a distractor is rewarded, they were faster at learning how to suppress it.***

Thus, Study 3 provided tentative support for the hypothesis that attentional capture is also partly a motivational phenomenon. Yet, as Study 3 was not designed to implement a between-subjects comparison, the results are exploratory and need further confirmation by additional studies. Study 3 also provided an insight concerning the manipulation of incentives in the additional singleton task: If the additional singleton task with a constant target (cf. Study 2) is used, manipulating reward block-wise can be an option if sequence effects are taken into account. If, however, the difficult version of the additional singleton task with a variable target (cf. Study 3) is used, a trial-wise manipulation of incentives seems to be the best option.

Like Study 1, Study 3 showed that phasic pupil dilation is increased in incentive conditions, but only in the later blocks of the experiment. In contrast to Study 1, tonic pupil dilation was measured, too, because incentives were manipulated block-wise and thus also sustained changes in mental effort could be detected. The data pattern of tonic pupil dilation was consistent with the hypothesis only in the neutral start group. Among participants of this group, pupil dilation increased in reward blocks and decreased in neutral blocks. In the reward start group, however, tonic pupil dilation decreased linearly over the course of the experiment. This provides evidence that the two sequence groups not only differed in their visual search performance, but also in their investment of mental effort.

Study 4 was aimed at challenging the emotional attentional capture account, which states that emotional faces can be preattentively processed and therefore capture attention automatically (Hodsoll et al., 2011; Öhman et al., 2001). This account is based on the reasoning of evolutionary psychology: the fast and automatic detection of possible threats is thought to be a prerequisite for survival in a dangerous environment (Öhman & Mineka, 2001). A closer look at studies on the emotional capture account revealed three methodological weaknesses. Firstly, many studies used task goals that were associated with emotional content, so that any emotional capture effect might actually have been a top-down effect. Secondly, some studies did not test whether search was parallel in the task at hand, so that they did not prove that a distraction effect is a capture effect. Thirdly, the problem of visual confounds was not addressed in all of the studies (see Chapter 2.3). Considering these aspects, the empirical evidence for emotional attentional capture is actually weak. Therefore, Study 4 tested the hypothesis that emotional faces do not capture attention, but distract search if the task goal requires the focal processing of the faces. This means that a distraction effect by emotional stimuli is dependent on the top-down task set. To explore this issue, an emotional variant of the additional singleton task was used, in which the singleton was not a discrepant feature, but a discrepant emotional face among neutral faces. Study 4 compared two task goals in a between-subjects design: One group of participants searched for a singleton color dot, the other group searched for the male face among female faces. Search was parallel in the former and serial in the latter group. In contrast to the emotional capture account, emotional faces did not delay search time or any other indicator of attentional processing in the group that searched for the color singleton. In the other group, however, search time was delayed by upright emotional faces. Further analysis also proved that the emotional faces did not affect indicators of attentional capture in this task. Instead, they drew additional, later fixations, during focal processing of the faces. Thus, Study 4 delivered convincing evidence that ***an emotional face does not capture attention preattentively, but interferes with attentional processing if faces need to be identified and search is serial***.

8.2 Limitations

The Discussion sections of Studies 1 - 4 already pointed out limitations that were specific for the respective study. This chapter discusses some limitations that are shared by all or most of the studies.

The first limitation involves the measurement of goal importance and therefore the validity of the goal importance manipulation. Participants were not asked to rate how important they regarded the goal of finding the target quickly and correctly, or the related goal of suppressing the CSD. Instead, the studies measured transient and sustained mental effort with pupil dilation. As expected, phasic pupil dilation was increased in incentive trials and in incentive blocks, thus strongly suggesting that participants invested more mental effort in the task when they expected a reward or loss. It was assumed that mental effort increased because participants regarded the task goal as more important in reward conditions, but this subjective judgment was not measured. Wykowska et al. (2013) measured participants' motivation during the task both via pupil dilation (Exp. 1) and via a visual analog scale (Exp. 2) and found consistent results in both studies. As the authors did not employ both pupil dilation and self-report measures in one experiment, they could not report a correlation between the two variables. The consistent correlation of pupil dilation and self-reported motivation with the dependent variable, however, at least suggest a correlation between pupil dilation and self-reported motivation. An interesting recent study (Miranda & Palmer, 2014) reported that rewarding participants for correct and fast responses with bonus points during an additional singleton task increased intrinsic motivation for the task (measured with a self-report questionnaire), in comparison to a control group. In contrast to Study 1, participants could only earn bonus points and could not lose. Another study showed that bonus points have equivalent effects compared to monetary rewards (Shen & Chun, 2011). These results strongly suggest that the payoff scheme employed in Study 1 and 3 was motivating, but a direct proof is missing.

In future studies, motivation of participants should be measured both implicitly with pupil dilation and more directly with self-report measures. This procedure would also allow testing an alternative interpretation of pupil dilation in Study 1 and Study 3. The similar data patterns for phasic pupil dilation and performance indicators strongly suggest that phasic pupil dilation measured mental effort, yet the lack of correlation indicates that pupil size was also determined by other mental processes. A study by Bradley, Miccoli, Escrig, and Lang (2008) reported that phasic pupil dilation increased when participants viewed positive or negative emotional pictures, compared to when they viewed pictures with neutral content. Based on this result, the authors argued that pupil dilation is also an indicator of emotional arousal. Consequently, one could propose that pupil dilation in Study 1 and Study 3 was not only determined by mental effort, but also by the emotional arousal due to the anticipation of rewards or losses. Yet, the results of Bradley et al. (2008) also show that emotional arousal only leads to a significant increase in pupil dilation after two seconds of picture viewing. Therefore, it seems questionable, though not impossible, that emotional arousal affected pupil dilation during the markedly shorter trials (less than 1 second) of Study 1 and Study 3. This issue has not yet been addressed by any of the studies analyzing reward and mental effort effects on pupil dilation (Bijleveld et al., 2009; Heitz et al., 2008; Wykowska et al., 2013). To test whether the cueing of reward also leads to emotional arousal, future studies should measure pupil dilation and both motivation and emotional arousal via self-report.

The study by Miranda and Palmer (2014) is also interesting for another reason: It shows that incentives that are embedded in a game-like environment and that are based either on bonus points or on monetary rewards are an effective and convenient way of increasing individuals' intrinsic motivation for an experimental task, especially for those tasks which are typically regarded as boring by participants (visual search tasks with several hundred of trials might be among them).

A second, related limitation concerns the manipulation of goal importance in the studies. Participants in Study 3 started with 0.00 € and only lost money if they made an error. In Study 1, participants started with 7.00 € and lost in loss trial whenever the response latency was slower than the adaptive time limit or when they made an error. On average, participants won 7.80 € in Study 1 and 4.54 € in Study 3. Thus, the amount of monetary reward increased markedly for participants in Study 3, whereas it did not change that much for participants in Study 1. The payoff scheme in Study 1 might have induced a prevention focus and the payoff scheme in Study 3 might have induced a promotion focus (Higgins, 1997). The comparison of error rates in Study 1 and in Study 3, however, does not support this assumption. If participants in Study 1 had adopted a prevention focus, one would expect a decreased error rate in Study 1 compared to that of Study 3. Yet, the error rate was actually higher in the former than in the latter. Nevertheless, possible differential effects of these outcome foci on motivation and on top-down control were not explicitly addressed in this doctoral thesis and would surely be an interesting topic for further research.

A third and fourth limitation regards two technological eye tracking procedures. In Studies 1, 3 and 4, a system-controlled calibration procedure was applied. During this kind of calibration procedure, the participant has to fixate several points on the computer screen one after another. Whereas the fixation on the first point is verified by the experimenter, the software checks whether the following fixations are valid. This is the standard calibration procedure integrated into the software of SensoMotoric Instruments and the other leading eye tracking manufacturers (Nyström, Andersson, Holmqvist, & van de Weijer, 2013). Nyström et al. (2013) examined the effect of different calibration procedures on the precision and accuracy of eye tracking data. They compared the system-controlled calibration to the participant-controlled and the operator-controlled calibration procedure. In a participant-controlled calibration procedure, the participant herself decides if and when she thinks that she properly fixates on the calibration stimulus. In an operator-controlled calibration procedure, the experimenter accepts a fixation as valid if she thinks that the participant is fixating on the target item properly. The results of the study showed that the participant-controlled calibration procedure significantly enhanced the precision and the accuracy of the eye tracking data in comparison to the other two procedures. This procedure would certainly have been a way to improve the eye tracking data quality, yet the study by Nyström et al. (2013) had not been published by the time the experiments had been programmed.

Furthermore, as was outlined in Chapter 2.1.2, the eye tracking manufacturers differ in the event detection mechanisms that they integrate into their software. In the studies of this doctoral thesis, event detection was conducted with the algorithm of the software BeGaze 3.4 by SensoMotoric Instruments, which allows the experimenter to set some event detection criteria. Two of those criteria were set to the same levels in all of the studies (minimum fixation duration: 50 ms, minimal saccadic velocity: 30°). The comparison between the eye movement data of Studies 1, 3 and 4 and the studies in the literature might be complicated by the fact that none of the other studies used an eye tracker by SensoMotoric Instruments. Therefore, those studies most probably used a different event detection algorithm. Theeuwes et al. (2003) and Geyer et al. (2008) reported that an eye movement was detected as a saccade if its velocity exceeded 35° or if its acceleration exceeded 9.500°/s^2. However, the setting of an acceleration criterion is not possible with the software of SensoMotoric Instruments. Holmqvist et al. (2011) argued that, depending on data quality and on the type of task, the eye movement variables that are computed by the different event detection algorithms can differ considerably. Unfortunately, the authors did not provide a quantitative analysis of these differences. Such a comparison would be an interesting topic of research, which could lead to recommendations on which event detection criteria to use in certain tasks. However, complete control over the process of event detection is not possible with any of the commercial software packages distributed by the eye tracking manufacturers. Nyström and Holmqvist (2010) have developed an event detection algorithm that can be adapted by the experimenter to the goals of the analysis and the characteristics of the eye tracking study. The application of such an algorithm to the raw gaze data, however, requires advanced programming skills and its usability is much lower compared to the commercial software packages.

A last limitation regards the sample size. Although it was comparable to the majority of studies examining attentional capture in the additional singleton task (Bacon & Egeth, 1994; Geyer et al., 2008; Leber & Egeth, 2006; Müller et al., 2009; Theeuwes, 1991, 1992, 2004; Theeuwes et al., 2003; Wienrich & Janczyk, 2011; Zehetleitner et al., 2012), it was not always sufficient to provide enough power to examine some hypotheses (correlation between visual working memory capacity and attentional capture in Study 2, comparison of sequence groups in Study 3, correlation between pupil size and mental effort in Study 1 and Study 3).

8.3 Challenges

8.3.1 The Difficulty of Measuring Attentional Capture with Eye Movements

It was a central aim of this doctoral thesis to examine visual search with indicators of overt and covert attention. Thus, eye movements and manual response were recorded in Studies 1 and 3. The additional singleton task that was used in these studies was initially developed to measure only covert attention. In the original version of this task (Theeuwes, 1991, 1992), eye movements were not necessary because the stimuli were placed very close to the fixation center, so that they could be perceived at least parafoveally.

If one wants to measure eye movements during this task, the stimulus eccentricity (i.e., the distance between the center of the screen and the center of the stimuli) has to be increased, so that participants need to make an eye movement to detect the target. In Studies 1 - 3, eccentricity was set to 11.4°, which is a bit larger than in the study by Theeuwes et al. (2003) and in the study by Geyer et al. (2008), but smaller than in the study by Müller et al. (2009) - a study that did not measure eye movements. Study 2, which used this eccentricity, combined with the easy version of the additional singleton task (constant target, homogeneous distractors), reported parallel search and attentional capture by the CSD. This result is in line with the other studies using the additional singleton task and measuring eye movements (Geyer et al., 2008; Theeuwes et al., 2003). However, search in the feature display group, which completed a more difficult version of the additional singleton task (heterogeneous distractors) was serial in Study 2. Although the distractor effect of the CSD was large, it was present for a much shorter period of time than in the study by Vatterott and Vecera (2012).

The decisive difference between the display of the feature display group in Study 2 and the displays used by Vatterott and Vecera (2012) and Theeuwes (1991) was the stimulus eccentricity, which was much larger in the former. This finding suggests that the easy version of the additional singleton task can be adapted for eye movement studies without changing the search process, while more difficult versions of the task (heterogeneous distractors in the feature display group in Study 2, variable target and distractors in Study 3) cannot be adapted to eye movement studies without inducing serial search. When search is serial, the global indicators of search cannot reveal anything about attentional capture because they measure both the preattentive and the serial stage of search. Eye movements, however, allow making inferences about attentional capture, even in studies in which search slopes are significantly different from zero. Yet, this kind of analysis has a severe disadvantage: In the serial search in Study 3, the hit rate of the first saccade on the target was much lower than in Study 1. Thus, the amount of trials that can be analyzed is heavily reduced if this variable is the only variable with which to make inferences about the search process. It follows from this conclusion that Study 1 was too easy to detect changes in overt and covert attentional capture, whereas Study 3 was not fully adequate to analyze overt and covert attentional capture because it was too difficult. Thus, the best method to analyze incentive effects on attentional capture, based on the results of this doctoral thesis, is to employ the easy version of the additional singleton task like in Study 1 and to change the distractor color after each block like in Study 2.

Furthermore, this argumentation suggests that the relationship between target eccentricity and parallel/serial search should be further investigated. Although some studies have already been conducted on eye movement recordings during the additional singleton task (Geyer et al., 2008; Theeuwes et al., 2003), more research with the aim to develop guidelines or standards for defining the "optimal" target eccentricity would be helpful to enable researchers to compare their results.

Finally, the problem of designing a task that is suitable for the monitoring of eye movements, difficult enough to evoke stable distractor interference, but also easy enough to be done in parallel search, hints at the elusiveness of attentional capture. Although it is assumed to be a basic bottom-up driven process (Theeuwes, 2010), it is quite difficult to detect, especially if one wants to measure it with eye movements. Ironically, this thesis' aim to deliver a more fine-grained analysis of attentional capture by monitoring eye movements might have complicated its observation. According to the active vision account (Findlay & Gilchrist, 2003), the function of covert attention is to prepare eye movements. The fact that the measurement of overt attentional capture is difficult and requires specific characteristics of the display suggests that the covert attentional capture effect (cf. Theeuwes, 1992) should not be considered as evidence for the hypothesis that the early phase of visual search is completely under bottom-up control.

8.3.2 The Differentiation between Parallel and Serial Search

Since the publication of feature integration theory (Treisman & Gelade, 1980), many theories of visual search categorize search as parallel if it produces a flat search slope and as serial if it is associated with a steep search slope of at least 10 ms/item (Müller & Krummenacher, 2006b). This differentiation is central to the stimulus-driven capture account by Theeuwes (2010), which argues that only parallel search is completely unaffected by top-down control. Thus, to test this account, one has to use a task that solely induces parallel search or to only analyze those components of covert search that represent the parallel stage of search. As has been pointed out in the previous chapter, designing such a task can be quite a difficult endeavor, especially if several factors of the display are varied.

Wolfe (1998b) criticized the dichotomy between strictly parallel and strictly serial search tasks as “mythical” (p. 34). He argued that feature search tasks (i.e., the target is defined by one single feature, e.g., its color), which are usually considered to be parallel search tasks, can also lead to serial search if the distractors are heterogeneous or if the distractor features are similar to the target. This finding is based on the theory of visual search by Duncan and Humphreys (1989), which states that visual search becomes more difficult with increasing target - distractor or with decreasing distractor - distractor similarity. Wolfe (1998b) aggregated search slopes from 2.500 search tasks conducted in his lab, ranging from very simple pop-out search tasks to difficult conjunction search tasks, and observed an unimodal distribution of the search slope. Thus, he argued that parallel and serial searches should rather be conceptualized as two ends of a continuum ranging from very efficient (search slope = 0 ms/item) to very inefficient search tasks (search slope > 30 ms/item). Search tasks with a slope of about 5 - 10 ms/item are classified as being “quite efficient” (Wolfe, 1998a, 1998b). Although Wolfe does not advise to use the terms parallel and serial to classify certain tasks, he admits that flat search slopes (0 ms/item) indicate a very different kind of search than steep search slopes. Thus, it seems quite easy to define what a parallel search is, but very difficult to categorize a search as serial. According to Wolfe’s “categorization”, the search slopes in the feature display task in Study 2 were still quite efficient. This leads to the important question: What does this mean for the observation of attentional capture? Is “quite efficient” already “inefficient” in the sense that it is not parallel any more, or is it still close enough to parallel?

This thesis provides at least some suggestions on how to generate answers to this question: Firstly, eye movements should be monitored whenever possible because they allow an inference about preattentive search, even if slopes indicate inefficient search. Secondly, further research on the relationship between search slopes and eye movements, especially in the additional singleton task is necessary. Theeuwes (2010) did not address Wolfe’s criticism of the dichotomy between parallel and serial search tasks. In a study (Theeuwes, 2004) on the attentional window hypothesis, he regarded a search slope that differed significantly from zero (11 ms/item) as a marker of serial search.

8.4 Conclusion

This thesis made a first step in studying goal-directed visual search from a different angle than previous research. First, it was its aim to examine whether the cognitive deficit of attentional capture in a search task is, to some extent, a motivational deficit. Direct evidence for this hypothesis could not be provided by the data, for reasons discussed in the previous chapters. Nevertheless, by trying to design a study that is adequate to measure both attentional capture with eye movements and motivational influences, it was observed that attentional capture is quite a transient effect that is quickly reduced by distractor practice. These results confirm the dimension weighting account (Müller et al., 2009; Zehetleitner et al., 2012) and show that attentional capture can be controlled by (cognitive) intentions of the observer. Furthermore, the studies provide tentative evidence that incentives can decrease attentional capture if they are adequately implemented. Second, the hypothesis that emotional attentional capture is actually due to top-down processes was strongly confirmed by the analyses. These results call for a closer theoretical and methodological integration of the cognitive and the motivational perspective on visual search in order to arrive at more precise explanations of how visual search works in the real life.

9 REFERENCES

Allport, A. (1989). Visual attention. In M. I. Posner (Ed.), *Foundations of cognitive science.* (pp. 631-682). Cambridge, MA US: The MIT Press.

Anderson, B. A., & Yantis, S. (2012). Value-driven attentional and oculomotor capture during goal-directed, unconstrained viewing. *Attention, Perception, & Psychophysics, 74*(8), 1644-1653. doi: 10.3758/s13414-012-0348-2

Atkinson, J. W. (1964). *An introduction to motivation*. Oxford England: Van Nostrand.

Austin, J. T., & Vancouver, J. B. (1996). Goal constructs in psychology: Structure, process, and content. *Psychological Bulletin, 120*(3), 338-375. doi: 10.1037/0033-2909.120.3.338

Bacon, W., & Egeth, H. (1994). Overriding stimulus-driven attentional capture. *Attention, Perception, & Psychophysics, 55*(5), 485-496. doi: 10.3758/bf03205306

Baldauf, D., & Deubel, H. (2008). Properties of attentional selection during the preparation of sequential saccades. *Experimental Brain Research, 184*(3), 411-425. doi: 10.1007/s00221-007-1114-x

Beatty, J. (1982). Task-evoked pupillary responses, processing load, and the structure of processing resources. *Psychological Bulletin, 91*(2), 276-292. doi: 10.1037/0033-2909.91.2.276

Beatty, J., & Kahneman, D. (1966). Pupillary changes in two memory tasks. *Psychonomic Science, 5*(10), 371-372.

Becker, D. V., Anderson, U. S., Mortensen, C. R., Neufeld, S. L., & Neel, R. (2011). The face in the crowd effect unconfounded: Happy faces, not angry faces, are more efficiently detected in single- and multiple-target visual search tasks. *Journal of Experimental Psychology: General, 140*(4), 637-659. doi: 10.1037/a002406010.1037/a0024060.supp (Supplemental)

BeGaze 3.4 [Software]. (2014). Teltow, Germany: SensoMotoric Instruments.

Belopolsky, A. V., Devue, C., & Theeuwes, J. (2011). Angry faces hold the eyes. *Visual Cognition, 19*(1), 27-36. doi: 10.1080/13506285.2010.536186

Belopolsky, A. V., & Theeuwes, J. (2010). No capture outside the attentional window. *Vision Research, 50*(23), 2543-2550. doi: 10.1016/j.visres.2010.08.023

Belopolsky, A. V., Zwaan, L., Theeuwes, J., & Kramer, A. F. (2007). The size of an attentional window modulates attentional capture by color singletons. *Psychonomic Bulletin & Review, 14*(5), 934-938. doi: 10.3758/bf03194124

Bijleveld, E., Custers, R., & Aarts, H. (2009). The unconscious eye opener: Pupil dilation reveals strategic recruitment of resources upon presentation of subliminal reward cues. *Psychological Science, 20*(11), 1313-1315. doi: 10.1111/j.1467-9280.2009.02443.x

Bodenhausen, G. V., & Hugenberg, K. (2009). Attention, perception, and social cognition. In F. Strack & J. Förster (Eds.), *Social cognition: The basis of human interaction.* (pp. 1-22). New York, NY US: Psychology Press.

Botvinick, M. M., Cohen, J. D., & Carter, C. S. (2004). Conflict monitoring and anterior cingulate cortex: An update. *Trends in Cognitive Sciences, 8*(12), 539-546. doi: 10.1016/j.tics.2004.10.003

Bradley, M. M., Miccoli, L., Escrig, M. A., & Lang, P. J. (2008). The pupil as a measure of emotional arousal and autonomic activation. *Psychophysiology, 45*(4), 602-607.

Braver, T. S. (2012). The variable nature of cognitive control: A dual mechanisms framework. *Trends in Cognitive Sciences, 16*(2), 106-113.

Braver, T. S., Krug, M. K., Chiew, K. S., Kool, W., Westbrook, J. A., Clement, N. J., . . . Somerville, L. H. (2014). Mechanisms of motivation-cognition interaction: challenges and opportunities. *Cognitive, Affective & Behavioral Neuroscience, 14*(2), 443-472. doi: 10.3758/s13415-014-0300-0

Brehm, J. W., & Self, E. A. (1989). The intensity of motivation. *Annual Review of Psychology, 40*, 109-131.

Broadbent, D. E. (1958). *Perception and communication*. Elmsford, NY US: Pergamon Press.

Calvo, M. G., & Nummenmaa, L. (2008). Detection of emotional faces: salient physical features guide effective visual search. *Journal of Experimental Psychology: General, 137*(3), 471-494. doi: 10.1037/a0012771

Calvo, M. G., Nummenmaa, L., & Avero, P. (2008). Visual search of emotional faces: Eye-movement assessment of component processes. *Experimental Psychology, 55*(6), 359-370. doi: 10.1027/1618-3169.55.6.359

Carver, C. S., & Scheier, M. F. (1998). *On the self-regulation of behavior*. New York, NY, US: Cambridge University Press.

Chiew, K. S., & Braver, T. S. (2011). Positive emotion, reward, and cognitive control: emotional versus motivational influences. *Frontiers in Psychology, 2*, 1-10. doi: 10.3389/fpsyg.2011.00279

Chiew, K. S., & Braver, T. S. (2013). Temporal dynamics of motivation-cognitive control interactions revealed by high-resolution pupillometry. *Frontiers in Psychology, 4*, 1-15. doi: 10.3389/fpsyg.2013.00015

Corbetta, M., & Shulman, G. L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nature Review Neuroscience, 3*(3), 201-215.

Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences, 24*(1), 87-185. doi: 10.1017/s0140525x01003922

Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience, 18*, 193-222.

Deubel, H., & Schneider, W. X. (1996). Saccade target selection and object recognition: Evidence for a common attentional mechanism. *Vision Research, 36*(12), 1827-1837. doi: 10.1016/0042-6989(95)00294-4

Devue, C., Belopolsky, A. V., & Theeuwes, J. (2012). Oculomotor Guidance and Capture by Irrelevant Faces. *PLoS ONE, 7*(4), 1-6. doi: 10.1371/journal.pone.0034598

DeWall, C. N., Baumeister, R. F., Mead, N. L., & Vohs, K. D. (2010). How leaders self-regulate their task performance: Evidence that power promotes diligence, depletion, and disdain. *Journal of Personality and Social Psychology*, 47-65. doi: 10.1037/a0020932

Driver, J. (2001). A selective review of selective attention research from the past century. *British Journal of Psychology, 92*(1), 53-78. doi: 10.1348/000712601162103

Duchowski, A. T. (2007). *Eye Tracking Methodology: Theory and Practice* (2 ed.). London: Springer.

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

Eastwood, J. D., Smilek, D., & Merikle, P. M. (2001). Differential attentional guidance by unattended faces expressing positive and negative emotion. *Perception & Psychophysics, 63*(6), 1004-1013.

Egeth, H. E., & Yantis, S. (1997). Visual attention: Control, representation, and time course. *Annual Review of Psychology, 48*, 269-297. doi: 10.1146/annurev.psych.48.1.269

Engelmann, J. B., & Pessoa, L. (2007). Motivation sharpens exogenous spatial attention. *Emotion, 7*(3), 668-674. doi: 10.1037/1528-3542.7.3.668

Farah, M. J., Tanaka, J. W., & Drain, H. M. (1995). What causes the face inversion effect? *Journal of Experimental Psychology: Human Perception and Performance, 21*(3), 628-634.

Findlay, J. M., & Gilchrist, I. D. (2003). *Active vision: The psychology of looking and seeing*. New York, NY US: Oxford University Press.

Findlay, J. M., & Walker, R. (1999). A model of saccade generation based on parallel processing and competitive inhibition. *Behavioral and Brain Sciences, 22*(4), 661-721.

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*(4), 1030-1044. doi: 10.1037/0096-1523.18.4.1030

Fox, E., Derakshan, N., & Standage, H. (2011). The assessment of human attention. In K. C. Klauer, A. Voss & C. Stahl (Eds.), *Cognitive methods in social psychology.* (pp. 15-47). New York, NY, US: Guilford Press.

Fox, E., Lester, V., Russo, R., Bowles, R. J., Pichler, A., & Dutton, K. (2000). Facial expressions of emotion: Are angry faces detected more efficiently? *Cognition and Emotion, 14*(1), 61-92.

Fox, E., Russo, R., Bowles, R., & Dutton, K. (2001). Do threatening stimuli draw or hold visual attention in subclinical anxiety? *Journal of Experimental Psychology: General, 130*(4), 681-700. doi: 10.1037//0096-3445.130.4.681

Fox, E., Russo, R., & Dutton, K. (2002). Attentional bias for threat: Evidence for delayed disengagement from emotional faces. *Cognition and Emotion, 16*(3), 355-379. doi: 10.1080/02699930143000527

Fröber, K., & Dreisbach, G. (2014). The differential influences of positive affect, random reward, and performance-contingent reward on cognitive control. *Cognitive, Affective & Behavioral Neuroscience*, 530-547. doi: 10.3758/s13415-014-0259-x

Fukuda, K., & Vogel, E. K. (2009). Human variation in overriding attentional capture. *The Journal of Neuroscience, 29*(27), 8726-8733. doi: 10.1523/jneurosci.2145-09.2009

Fukuda, K., & Vogel, E. K. (2011). Individual differences in recovery time from attentional capture. *Psychological Science, 22*(3), 361-368. doi: 10.1177/0956797611398493

Geyer, T., Müller, H. J., & Krummenacher, J. (2008). Expectancies modulate attentional capture by salient color singletons. *Vision Research, 48*(11), 1315-1326. doi: 10.1016/j.visres.2008.02.006

Godijn, R., & Theeuwes, J. (2002). Programming of endogenous and exogenous saccades: Evidence for a competitive integration model. *Journal of Experimental Psychology: Human Perception and Performance, 28*(5), 1039-1054. doi: 10.1037/0096-1523.28.5.1039

Goeleven, E., De Raedt, R., Leyman, L., & Verschuere, B. (2008). The Karolinska Directed Emotional Faces: A validation study. *Cognition and Emotion, 22*(6), 1094-1118.

Goldinger, S. D., He, Y., & Papesh, M. H. (2009). Deficits in cross-race face learning: Insights from eye movements and pupillometry. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 35*(5), 1105-1122. doi: 10.1037/a0016548

Gollwitzer, P. M. (1990). Action phases and mind-sets. In E. T. Higgins & R. M. Sorrentino (Eds.), *Handbook of motivation and cognition: Foundations of social behavior, Vol. 2.* (pp. 53-92). New York, NY, US: Guilford Press.

Gollwitzer, P. M., Kappes, H. B., & Oettingen, G. (2012). Needs and incentives as sources of goals. In H. Aarts & A. J. Elliot (Eds.), *Goal-directed behavior.* (pp. 115-150). New York, NY, US: Psychology Press.

Hansen, C. H., & Hansen, R. D. (1988). Finding the face in the crowd: An anger superiority effect. *Journal of Personality and Social Psychology, 54*(6), 917-924. doi: 10.1037/0022-3514.54.6.917

Heckhausen, H., & Gollwitzer, P. M. (1987). Thought contents and cognitive functioning in motivational versus volitional states of mind. *Motivation and Emotion, 11*(2), 101-120.

Heitz, R. P., Schrock, J. C., Payne, T. W., & Engle, R. W. (2008). Effects of incentive on working memory capacity: Behavioral and pupillometric data. *Psychophysiology, 45*(1), 119-129.

Hess, E. H., & Polt, J. M. (1964). Pupil size in relation to mental activity during simple problem-solving. *Science, 143*(Whole No. 3611), 1190-1192.

Higgins, E. T. (1997). Beyond pleasure and pain. *American Psychologist, 52*(12), 1280-1300.

Hills, P. J., & Pake, J. M. (2013). Eye-tracking the own-race bias in face recognition: Revealing the perceptual and socio-cognitive mechanisms. *Cognition, 129*(3), 586-597.

Hodsoll, S., Viding, E., & Lavie, N. (2011). Attentional capture by irrelevant emotional distractor faces. *Emotion, 11*(2), 346-353. doi: 10.1037/a0022771

Hoffman, J. E., & Subramaniam, B. (1995). The role of visual attention in saccadic eye movements. *Perception & Psychophysics, 57*(6), 787-795.

Holmqvist, K., Nyström, M., Andersson, R., Dewhurst, R., Jarodzka, H., & van de Weijer, J. (2011). *Eye tracking : a comprehensive guide to methods and measures*. Oxford [u.a.]: Oxford University Press.

Hooge, I. T. C., & Erkelens, C. J. (1999). Peripheral vision and oculomotor control during visual search. *Vision Research, 39*(8), 1567-1575.

Horstmann, G. (2005). Attentional Capture by an Unannounced Color Singleton Depends on Expectation Discrepancy. *Journal of Experimental Psychology: Human Perception and Performance, 31*(5), 1039-1060. doi: 10.1037/0096-1523.31.5.1039

Horstmann, G., Becker, S. I., Bergmann, S. I., & Burghaus, L. (2010). A reversal of the search asymmetry favouring negative schematic faces. *Visual Cognition, 18*(7), 981-1016.

Huang, S.-L., Chang, Y.-C., & Chen, Y.-J. (2011). Task-irrelevant angry faces capture attention in visual search while modulated by resources. *Emotion, 11*(3), 544-552. doi: 10.1037/a0022763

Hunt, A. R., Cooper, R. M., Hungr, C., & Kingstone, A. (2007). The effect of emotional faces on eye movements and attention. *Visual Cognition, 15*(5), 513-531. doi: 10.1080/13506280600843346

Hyönä, J., Tommola, J., & Alaja, A.-M. (1995). Pupil dilation as a measure of processing load in simultaneous interpretation and other language tasks. *The Quarterly Journal of Experimental Psychology A: Human Experimental Psychology, 48A*(3), 598-612.

iView X™ Hi-Speed 1250 Hz [Apparatus and software]. (2010). Teltow, Germany: SensoMotoric Instruments.

Just, M. A., & Carpenter, P. A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review, 87*(4), 329-354.

Juth, P., Lundqvist, D., Karlsson, A., & Öhman, A. (2005). Looking for Foes and Friends: Perceptual and Emotional Factors When Finding a Face in the Crowd. *Emotion, 5*(4), 379-395.

Kahneman, D., & Tversky, A. (1984). Choices, values, and frames. *American Psychologist, 39*(4), 341-350.

Kiss, M., Driver, J., & Eimer, M. (2009). Reward priority of visual target singletons modulates event-related potential signatures of attentional selection. *Psychological Science, 20*(2), 245-251. doi: 10.1111/j.1467-9280.2009.02281.x

Kool, W., McGuire, J. T., Rosen, Z. B., & Botvinick, M. M. (2010). Decision making and the avoidance of cognitive demand. *Journal of Experimental Psychology: General, 139*(4), 665-682.

Kruglanski, A. W. (1996). Goals as knowledge structures. In P. M. Gollwitzer & J. A. Bargh (Eds.), *The psychology of action: Linking cognition and motivation to behavior.* (pp. 599-618). New York, NY US: Guilford Press.

Kruglanski, A. W., Bélanger, J. J., Chen, X., Köpetz, C., Pierro, A., & Mannetti, L. (2012). The energetics of motivated cognition: A force-field analysis. *Psychological Review, 119*(1), 1-20. doi: 10.1037/a0025488

Krummenacher, J., & Müller, H. J. (2012). Dynamic weighting of feature dimensions in visual search: behavioral and psychophysiological evidence. *Frontiers in Psychology, 3*, 1-12. doi: 10.3389/fpsyg.2012.00221

Krummenacher, J., Müller, H. J., & Heller, D. (2003). Dimensionsbasierte Aufmerksamkeit bei visueller Suche. *Psychologische Rundschau, 54*(1), 24-34. doi: 10.1026//0033-3042.54.1.24

Kuchinke, L., Võ, M. L. H., Hofmann, M., & Jacobs, A. M. (2007). Pupillary responses during lexical decisions vary with word frequency but not emotional valence. *International Journal of Psychophysiology, 65*(2), 132-140. doi: 10.1016/j.ijpsycho.2007.04.004

Kuhl, J. (Ed.). (1994). *Volition and personality : action versus state orientation*. Seattle ; Göttingen [u.a.]: Hogrefe.

Laeng, B., Sirois, S., & Gredebäck, G. (2012). Pupillometry: A window to the preconscious? *Perspectives on Psychological Science, 7*(1), 18-27.

Langner, O., Dotsch, R., Bijlstra, G., Wigboldus, D. H. J., Hawk, S. T., & van Knippenberg, A. (2010). Presentation and validation of the Radboud Faces Database. *Cognition and Emotion, 24*(8), 1377-1388. doi: 10.1080/02699930903485076

Lavie, N., & de Fockert, J. (2006). Frontal control of attentional capture in visual search. *Visual Cognition, 14*(4-8), 863-876. doi: 10.1080/13506280500195953

Leber, A. B., & Egeth, H. E. (2006). It's under control: Top-down search strategies can override attentional capture. *Psychonomic Bulletin & Review, 13*(1), 132-138. doi: 10.3758/bf03193824

Leber, A. B., Kawahara, J.-I., & Gabari, Y. (2009). Long-term abstract learning of attentional set. *Journal of Experimental Psychology: Human Perception and Performance, 35*(5), 1385-1397. doi: 10.1037/a0016470

Lewin, K., Heider, F., & Heider, G. M. (1936). *Principles of topological psychology*. New York, NY, US: McGraw-Hill.

Locke, H. S., & Braver, T. S. (2008). Motivational influences on cognitive control: Behavior, brain activation, and individual differences. *Cognitive, Affective & Behavioral Neuroscience, 8*(1), 99-112. doi: 10.3758/cabn.8.1.99

Lu, S., & Han, S. (2009). Attentional capture is contingent on the interaction between task demand and stimulus salience. *Attention, Perception, & Psychophysics, 71*(5), 1015-1026. doi: 10.3758/app.71.5.1015

Maslow, A. H. (1943). A theory of human motivation. *Psychological Review, 50*(4), 370-396.

Maurer, D., Le Grand, R., & Mondloch, C. J. (2002). The many faces of configural processing. *Trends in Cognitive Sciences, 6*(6), 255-260.

McClelland, D. C. (1985). How motives, skills, and values determine what people do. *American Psychologist, 40*(7), 812-825.

Miller, E. K., & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience, 24*, 167-202.

Miranda, A., & Palmer, E. (2014). Intrinsic motivation and attentional capture from gamelike features in a visual search task. *Behavior Research Methods, 46*(1), 159-172. doi: 10.3758/s13428-013-0357-7

Müller, H. J., Geyer, T., Zehetleitner, M., & Krummenacher, J. (2009). Attentional capture by salient color singleton distractors is modulated by top-down dimensional set. *Journal of*

Experimental Psychology: Human Perception and Performance, 35(1), 1-16. doi: 10.1037/0096-1523.35.1.1

Müller, H. J., Heller, D., & Ziegler, J. (1995). Visual search for singleton feature targets within and across feature dimensions. *Perception & Psychophysics, 57*(1), 1-17. doi: 10.3758/BF03211845

Müller, H. J., & Krummenacher, J. (2006a). Locus of dimension weighting: Preattentive or postselective? *Visual Cognition, 14*(4-8), 490-513. doi: 10.1080/13506280500194154

Müller, H. J., & Krummenacher, J. (2006b). Visual search and selective attention. *Visual Cognition, 14*(4-8), 389-410. doi: 10.1080/13506280500527676

Muraven, M., & Slessareva, E. (2003). Mechanism of self-control failure: Motivation and limited resources. *Personality and Social Psychology Bulletin, 29*(7), 894-906.

Nyström, M., Andersson, R., Holmqvist, K., & van de Weijer, J. (2013). The influence of calibration method and eye physiology on eyetracking data quality. *Behavior Research Methods, 45*(1), 272-288. doi: 10.3758/s13428-012-0247-4

Nyström, M., & Holmqvist, K. (2010). An adaptive algorithm for fixation, saccade, and glissade detection in eyetracking data. *Behavior Research Methods, 42*(1), 188-204. doi: 10.3758/BRM.42.1.188

Öhman, A., Lundqvist, D., & Esteves, F. (2001). The face in the crowd revisited: A threat advantage with schematic stimuli. *Journal of Personality and Social Psychology, 80*(3), 381-396. doi: 10.1037/0022-3514.80.3.381

Öhman, A., & Mineka, S. (2001). Fears, phobias, and preparedness: Toward an evolved module of fear and fear learning. *Psychological Review, 108*(3), 483-522. doi: 10.1037//0033-295x.108.3.483

Pashler, H. E. (Ed.). (1998). *Attention*. Hove: Psychology Press.

Porter, G., Troscianko, T., & Gilchrist, I. D. (2007). Effort during visual search and counting: Insights from pupillometry. *The Quarterly Journal of Experimental Psychology, 60*(2), 211-229. doi: 10.1080/17470210600673818

Poynton, C. (2003). *Digital Video and HDTV: Algorithms and Interfaces*. San Francisco: Morgan Kaufmann.

Purcell, D. G., Stewart, A. L., & Skov, R. B. (1996). It takes a confounded face to pop out of a crowd. *Perception, 25*(9), 1091-1108.

Rayner, K. (2009). Eye movements and attention in reading, scene perception, and visual search. *The Quarterly Journal of Experimental Psychology, 62*(8), 1457-1506. doi: 10.1080/17470210902816461

RB-740 response pad [Apparatus]. (2013). San Pedro, USA: Cedrus Corporation.

Rizzolatti, G., Riggio, L., Dascola, I., & Umiltá, C. (1987). Reorienting attention across the horizontal and vertical meridians: Evidence in favor of a premotor theory of attention. *Neuropsychologia, 25*(1-A), 31-40.

Rothermund, K. (2011). Counter-regulation and control-dependency: Affective processing biases in the service of action regulation. *Social Psychology, 42*(1), 56-66. doi: 10.1027/1864-9335/a000043

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist, 55*(1), 68-78.

Savine, A. C., Beck, S. M., Edwards, B. G., Chiew, K. S., & Braver, T. S. (2010). Enhancement of cognitive control by approach and avoidance motivational states. *Cognition and Emotion, 24*(2), 338-356. doi: 10.1080/02699930903381564

Savine, A. C., & Braver, T. S. (2010). Motivated cognitive control: Reward incentives modulate preparatory neural activity during task-switching. *The Journal of Neuroscience, 30*(31), 10294-10305.

Schreij, D., Owens, C., & Theeuwes, J. (2008). Abrupt onsets capture attention independent of top-down control settings. *Perception & Psychophysics, 70*(2), 208-218. doi: 10.3758/pp.70.2.208

Scialfa, C. T., & Joffe, K. M. (1998). Response times and eye movements in feature and conjunction search as a function of target eccentricity. *Perception & Psychophysics, 60*(6), 1067-1082.

Shen, Y. J., & Chun, M. M. (2011). Increases in rewards promote flexible behavior. *Attention, Perception, & Psychophysics, 73*(3), 938-952. doi: 10.3758/s13414-010-0065-7

Shepherd, M., Findlay, J. M., & Hockey, R. J. (1986). The relationship between eye movements and spatial attention. *The Quarterly Journal of Experimental Psychology A: Human Experimental Psychology, 38*(3-A), 475-491.

Siebold, A., van Zoest, W., & Donk, M. (2011). Oculomotor evidence for top-down control following the initial saccade. *PLoS ONE, 6*(9), 1-10. doi: 10.1371/journal.pone.0023552

Tabachnick, B. G., & Fidell, L. S. (2014). *Using multivariate statistics* (6. ed., Pearson new international edition ed.). Harlow: Pearson Education Ltd.

Theeuwes, J. (1991). Cross-dimensional perceptual selectivity. *Perception & Psychophysics, 50*(2), 184-193.

Theeuwes, J. (1992). Perceptual selectivity for color and form. *Perception & Psychophysics, 51*(6), 599-606.

Theeuwes, J. (2004). Top-down search strategies cannot override attentional capture. *Psychonomic Bulletin & Review, 11*(1), 65-70. doi: 10.3758/bf03206462

Theeuwes, J. (2010). Top–down and bottom–up control of visual selection. *Acta Psychologica, 135*(2), 77-99. doi: 10.1016/j.actpsy.2010.02.006

Theeuwes, J., DeVries, G.-J., & Godijn, R. (2003). Attentional and oculomotor capture with static singletons. *Perception & Psychophysics, 65*(5), 735-746. doi: 10.3758/bf03194810

Trappenberg, T. P., Dorris, M. C., Munoz, D. P., & Klein, R. M. (2001). A model of saccade initiation based on the competitive integration of exogenous and endogenous signals in the superior colliculus. *Journal of Cognitive Neuroscience, 13*(2), 256-271.

Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology, 12*(1), 97-136. doi: 10.1016/0010-0285(80)90005-5

van Zoest, W., Donk, M., & Theeuwes, J. (2004). The Role of Stimulus-Driven and Goal-Driven Control in Saccadic Visual Selection. *Journal of Experimental Psychology: Human Perception and Performance, 30*(4), 746-759. doi: 10.1037/0096-1523.30.4.749

Vatterott, D. B., & Vecera, S. P. (2012). Experience-dependent attentional tuning of distractor rejection. *Psychonomic Bulletin & Review, 19*(5), 871-878. doi: 10.3758/s13423-012-0280-4

Veling, H., & Aarts, H. (2010). Cueing task goals and earning money: Relatively high monetary rewards reduce failures to act on goals in a Stroop task. *Motivation and Emotion, 34*(2), 184-190. doi: 10.1007/s11031-010-9160-2

Vogel, E. K., Woodman, G. F., & Luck, S. J. (2001). Storage of features, conjunctions, and objects in visual working memory. *Journal of Experimental Psychology: Human Perception and Performance, 27*(1), 92-114. doi: 10.1037/0096-1523.27.1.92

Vogt, J., De Houwer, J., & Crombez, G. (2011). Multiple goal management starts with attention: Goal prioritizing affects the allocation of spatial attention to goal-relevant events. *Experimental Psychology, 58*(1), 55-61. doi: 10.1027/1618-3169/a000066

Vogt, J., De Houwer, J., Moors, A., Van Damme, S., & Crombez, G. (2010). The automatic orienting of attention to goal-relevant stimuli. *Acta Psychologica, 134*(1), 61-69. doi: 10.1016/j.actpsy.2009.12.006

Voss, A. (2004). *Motivierte Wahrnehmung: selektive Aufmerksamkeit und entlastende Interpretationen bei der Aufnahme valenter Informationen.* (Doctoral dissertation, University of Trier). Retrieved from http://ubt.opus.hbz-nrw.de/ database.

Voss, A., Leonhart, R., & Stahl, C. (2007). How to make your own response boxes: A step-by-step guide for the construction of reliable and inexpensive parallel-port response pads from computer mice. *Behavior Research Methods, 39*(4), 797-801. doi: 10.3758/bf03192971

Wadlinger, H. A., & Isaacowitz, D. M. (2011). Fixing our focus: Training attention to regulate emotion. *Personality and Social Psychology Review, 15*(1), 75-102. doi: 10.1177/1088868310365565

Wentura, D., Müller, P., & Rothermund, K. (2013). Attentional capture by evaluative stimuli: Gain- and loss-connoting colors boost the additional-singleton effect. *Psychonomic Bulletin & Review*. doi: 10.3758/s13423-013-0531-z

Wienrich, C., & Janczyk, M. (2011). Absence of attentional capture in parallel search is possible: A failure to replicate attentional capture in a non-singleton target search task. *Attention, Perception, & Psychophysics, 73*(7), 2044-2052. doi: 10.3758/s13414-011-0183-x

Williams, D. E., Reingold, E. M., Moscovitch, M., & Behrmann, M. (1997). Patterns of eye movements during parallel and serial visual search tasks. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale, 51*(2), 151-164.

Williams, M. A., Moss, S. A., Bradshaw, J. L., & Mattingley, J. B. (2005). Look at me, I'm smiling: Visual search for threatening and nonthreatening facial expressions. *Visual Cognition, 12*(1), 29-50.

Wolfe, J. M. (1994). Guided search 2.0: A revised model of visual search. *Psychonomic Bulletin & Review, 1*(2), 202-238.

Wolfe, J. M. (1998a). Visual search. In H. Pashler (Ed.), *Attention.* (pp. 13-73). Hove, England: Psychology Press/Erlbaum (UK) Taylor & Francis.

Wolfe, J. M. (1998b). What can 1 million trials tell us about visual search? *Psychological Science, 9*(1), 33-39. doi: 10.1111/1467-9280.00006

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*(3), 419-433. doi: 10.1037/0096-1523.15.3.419

Wolfe, J. M., & Horowitz, T. S. (2004). What attributes guide the deployment of visual attention and how do they do it? *Nature Reviews Neuroscience, 5*(6), 495-501.

Wykowska, A., Anderl, C., Schubö, A., & Hommel, B. (2013). Motivation modulates visual attention: Evidence from pupillometry. *Frontiers in Psychology, 4*, 1-12. doi: 10.3389/fpsyg.2013.00059

Yantis, S. (2000). I: Control of visual attention. In S. Monsell & J. Driver (Eds.), *Control of cognitive processes: Attention and performance XVIII.* (pp. 71-208). Cambridge, MA US: The MIT Press.

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*(3), 661-676. doi: 10.1037/0096-1523.25.3.661

Yiend, J. (2010). The effects of emotion on attention: A review of attentional processing of emotional information. *Cognition and Emotion, 24*(1), 3-47. doi: 10.1080/02699930903205698

Zehetleitner, M., Goschy, H., & Müller, H. J. (2012). Top-down control of attention: It's gradual, practice-dependent, and hierarchically organized. *Journal of Experimental Psychology: Human Perception and Performance, 38*(4), 941-957. doi: 10.1037/a0027629

Zelinsky, G. J. (2008). A theory of eye movements during target acquisition. *Psychological Review, 115*(4), 787-835. doi: 10.1037/a0013118

Zelinsky, G. J., & Sheinberg, D. L. (1997). Eye movements during parallel–serial visual search. *Journal of Experimental Psychology: Human Perception and Performance, 23*(1), 244-262.

10 LIST OF TABLES

11 LIST OF FIGURES

12 LIST OF ABBREVIATIONS

CSD	color singleton distractor
DTG	dot task group
Exp.	experiment
fMRI	functional magnetic resonance imaging
MTG	male task group

DECLARATION

Erklärung gemäß § 8 Abs. 1 Buchst. b) und c) der Promotionsordnung der Fakultät für Verhaltens- und Empirische Kulturwissenschaften

Declaration in accordance to § 8 (1) b) and § 8 (1) c) of the doctoral degree regulation of Heidelberg University, Faculty of Behavioural and Cultural Studies

Promotionsausschuss der Fakultät für Verhaltens- und Empirische Kulturwissenschaften der Ruprecht-Karls-Universität Heidelberg
Doctoral Committee of the Faculty of Behavioural and Cultural Studies of Heidelberg University

Erklärung gemäß § 8 Abs. 1 Buchst. b) der Promotionsordnung der Universität Heidelberg für die Fakultät für Verhaltens- und Empirische Kulturwissenschaften
Declaration in accordance to § 8 (1) b) and § 8 (1) c) of the doctoral degree regulation of Heidelberg University, Faculty of Behavioural and Cultural Studies

Ich erkläre, dass ich die vorgelegte Dissertation selbstständig angefertigt, nur die angegebenen Hilfsmittel benutzt und die Zitate gekennzeichnet habe.
I declare that I have made the submitted dissertation independently, using only the specified tools and have correctly marked all quotations.

Erklärung gemäß § 8 Abs. 1 Buchst. c) der Promotionsordnung der Universität Heidelberg für die Fakultät für Verhaltens- und Empirische Kulturwissenschaften

Ich erkläre, dass ich die vorgelegte Dissertation in dieser oder einer anderen Form nicht anderweitig als Prüfungsarbeit verwendet oder einer anderen Fakultät als Dissertation vorgelegt habe.
I declare that I did not use the submitted dissertation in this or any other form as an examination paper until now and that I did not submit it in another faculty.

Vorname Nachname
First name Family name ______________________

Datum, Unterschrift
Date, Signature ______________________