Visual Search and Task-irrelevant Shape Information in Autism Spectrum Disorder

Mara Tribull

UCL

Thesis submitted for the Degree of Master of Philosophy in Computer Science

December 2014

Declaration

I, Mara Tribull, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Abstract

Visual processing in autism has become a popular topic of research in the past two decades. However, many findings in the reported visual processing abnormalities in autism remain contradictory (Greenaway & Plaisted-Grant, 2013; Simmons et al., 2009). Theories of autism divide into those that favor an explanation of symptoms based on differences in social cognition and orienting on the one hand, and on perception, including vision, on the other.

This thesis explores potential differences in visual attentional processing in autism spectrum disorder independently of social cognition factors. The focus lies on visual search as a task by which shifts of attention can be measured and inferences drawn as to underlying perceptual and cognitive processes. It reviews previous work on visual search in autism. The only recently published review on this topic lacked scrutiny (Kaldy et al., 2013). Reviewing the studies with respect to task designs and demographic features of the participants reveals a pattern of results consistent with a previously reported pattern (O'Riordan et al., 2001) but which is distinct from the common notion of "superior search abilities in autism in general". While children with autism are faster in conjunction and only inefficient feature searches, findings are mixed with regard to adults and feature search and are largely lacking for conjunction search in adults with autism.

This thesis also presents an experiment which makes use of the combination of eye-tracking techniques and a visual orientation-feature search task under interference from higher level shape recognition (Zhaoping & Guyader, 2007). Moreover, relevant studies addressing what may be called "task-irrelevant higher-level visual processing" in autism are reviewed.

Finally, a more complete characterization of the disorder in terms of its symptomatology with regard to visual attention and visual perception might lead to a better understanding of its causes.

Table of contents

Table of tables

Table of figures

Abbreviations

M: mean

SD: standard deviation

IQ: intelligence quotient

ANOVA: analysis of variance

Acknowledgments

I want to thank those who supported me during this thesis project.

My supervisor, Li Zhaoping, for giving me many opportunities to grow both as a researcher and as a person.

Miguel Angel Vadillo for his generosity in advising me during the tough times of writing up, for interesting discussions, and, most importantly, for believing in me.

Sarah White, for introducing me to testing people with autism. Thanks also to all participants of our study.

Adrian Hellekes, for lively discussions and for pointing out to me perspectives I would never have thought of; Valentina Rueda Castro, Johanna Finnemann, Lawrence York, and Emmanouil Protonotarios, for their comments, discussions, and for reminding me of the big picture when I was immersed in details; Renata Mangano, Ge XinXin—I was very lucky to get the chance to collaborate with them. Steffen Grünewälder, for many tea breaks and creating an enjoyable environment at the office. All my friends for their support. A big thanks to Annica Flemming for correcting my English.

Those who have always been there and reassured me throughout the past three years (and as long as I can remember): My parents.

Tirso for his love, reassurance, and endless patience. I would not have endured until now if it was not for him.

Preface

The present thesis was written under the premise that I am not an advocate or expert of any of the theories discussed, and that, as a result, there is no reason to confine the reader to a single viewpoint. Instead, my aim was to be critical and to take into account several different viewpoints on the same topic. Nevertheless, I am aware that I was inevitably influenced by some theoretical frameworks, just like all observations are theory-laden[1]. In particular, the view that one can distinguish three processing stages in vision—encoding (of low level visual features), selection (by attention), and decoding (by higher cognitive processing)—might not be the common view. It is part of a general theoretical framework promoted by Li Zhaoping. Another author whose work is predominant in the field of visual search is Jeremy Wolfe. His influence in my writing is probably already visible in the proportion of text dedicated to visual search tasks, such that the introductory chapter of the present thesis is dedicated fully to this task.

The order of presentation in Chapter 1 is arbitrary in principle but follows the general logic of "behavioral observations, theories about underlying mechanisms, biological evidence". This is followed by Chapter 2, which reviews previous work on visual search in children and adults with autism. Chapter 3 presents an experiment that was conducted in collaboration with Li Zhaoping and Sarah White and part of which was reported in a bachelor's thesis (Tribull, 2011). Chapter 4 provides a discussion of the results and limitations of this study as well as implications for future investigations. The present thesis closes with a few concluding remarks in Chapter 5.

---

[1] Kuhn, T.S., *The Structure of Scientific Revolutions*, 1962, Chicago: University of Chicago Press, reprinted,1996

Chapter 1 What has visual search told us about visual attention?

1.1 Visual search: General task design

Visual search, in its most general definition, is a task that requires the observer to find one or more targets within a display containing distractor items (Wolfe, 2014). The target and distractors can be shapes, textures, objects or even elements of a naturalistic visual stimulus. Visual search tasks are commonly used in connection with the study of visual attention. Visual attention is often seen as a filter process that reduces the flood of information coming in through our senses by letting only part of it pass through to higher cognitive processing. There must be some selection criteria that give priority to the important information over unimportant “background noise”. These criteria can change from one situation to another, as what is important in one situation may not be important in another. For instance, looking out for a green light when crossing a street may not be relevant in a psychological experiment where the aim of the task is to look out for a red flashing light whilst ignoring all green items.

The instruction “look out for …” guides attention to potential targets. This type of guidance is often referred to as “top-down” control over attention. The influences of instructions on attentional deployment are exemplified by Yarbus (1967). He presented observers with a picture of a scene containing several people in a room. Eye-movement patterns differed greatly according to the instructions for exploration of the scene within the same observer (e.g., “Give the ages of the people” vs. “Estimate material circumstances of the family”). In the selective attention literature the term “top-down” most often makes reference to the observer’s internal state, as opposed to automatic “bottom-up” control over attention, which depends on stimulus attributes that draw on the observer’s attention—even against her will in some circumstances. For example, in the additional singleton task the presence of a task-irrelevant unique item in the display significantly slows the search for the task-relevant unique item (Theeuwes, 1992). Theeuwes presented participants with uniformly colored items arranged in a circle. These could be all circles and a single diamond (singleton), or vice versa. Line segments of different orientations appeared within the shapes and observers had to determine the orientation of the segment within the target, which could be identified by its unique shape. In the distractor condition, one of the non-target items differed in color from all other items, making it a unique distractor item. Time to report the orientation within the shape target increased significantly in this condition.

According to Frith (2005) there are three explanatory levels on which the term “top-down” can be understood: One makes reference to the processing- and anatomical hierarchy of brain areas (see Box 1) and is often used by neuroscientists. Here, top-down control is exhibited by areas at higher levels in the hierarchy upon lower level areas. On the psychological level, top-down refers to the observer’s internal state, which can be influenced by task instructions, such as “look out for the dim red square”. This type of top-down control over attention is also sometimes called endogenous or goal-directed attention. On a third level, the phenomenological level, the term refers to the observer’s subjective impression of her voluntary deployment of attention. Bottom-up control over attention must always be understood in relation to the respective meaning of top-down control. Characteristics of top-down and bottom-up control of attention are opposed to each other. However, they must not be seen as a dichotomy. As Anderson (2011, p.2) points out:

> “This [dichotomies-] approach to attention research is harmful because instead of addressing the core phenomenal components we pursue a pseudo-question: is our dichotomy true?”

Most researchers seem to agree that both bottom-up and top-down control influence attention in our daily life, but that they can be distinguished in well-designed experiments. For the purpose of this thesis, types of top-down control will be distinguished where necessary: By “top-down in the brain” I refer to the usage of the term with reference to the anatomical hierarchy of brain areas. By “top-down in attention” I refer to its usage by psychologists, where it refers to factors that depend on the observer’s internal state (such as current task goals). Bottom-up, on the contrary, refers to factors that are dependent on stimulus characteristics and not on the observer’s goals, beliefs or intentions (Theeuwes, 2004).

BOX 1: Hierarchical visual processing

Most of what can be counted as “visual information processing” happens in anatomically and functionally distinguishable pathways. Within each pathway there is a sequential organization which is usually referred to as hierarchical processing (Kandel & Wurtz, 2000). Analysis of the connectivity patterns among visual areas in the monkey brain has led to a visual hierarchy model including 32 areas engaged in purely visual or visual-associative processing (Felleman & Van Essen, 1991). Visual input is analyzed in several parallel pathways starting as early as the retina (Livingstone & Hubel, 1988). From there signals are projected to the lateral geniculate nucleus of the thalamus and then to primary visual cortex (V1). Areas beyond V1 have been attributed to two major functionally separable parallel pathways, the dorsal stream, and the ventral stream (Mishkin, Ungerleider, & Macko, 1983; Goodale & Milner 1991). Starting from V1, the dorsal pathway continues to area V2 and from there to V3, MT (aka V5), MST, and further stations in the inferior parietal and superior temporal sulcal cortex (Ungerleider & Haxby 1994). The ventral pathway, in turn, includes areas V1, V2, V4, TEO and TE (Ungerleider & Haxby 1994). Areas along these pathways differ with respect to the dimensions of visual input that activate the areas or their participating cells, and with respect to their response latencies: The ventral pathway is thought to be mainly involved in the analysis of objects, whereas the dorsal stream has been associated with spatial vision (Mishkin et al., 1983) and the visual control of actions (Goodale & Milner, 1991). Areas along the dorsal stream respond to visual stimulation with shorter latencies compared to ventral stream areas (Schmolesky, Wang, Hanes, Thompson, Leutgeb, Shall et al., 1998).

The behavior of observers during visual search has been explained by theoretical models, some of which focus purely on bottom-up factors, others combine top-down and bottom-up factors to explain visual search behavior. Below I introduce some examples of now classic models and the criticisms they received, in particular with regard to the concept of a preattentive processing stage. After that, I briefly discuss findings about the underlying neural processes during visual search from neuroimaging and physiology studies.

1.1.1 The role of attention in visual search

I shall discuss briefly why we talk about attentional selection, or selective attention. We are never aware of *all* information that stimulates the two retinas. So, where does the majority of information get filtered out? And what factors determine which information is filtered out and which selected for further processing? It is widely assumed that attention mediates this selection. Similarly, the perhaps least disputed property of attention is its selective character (Itti, Rees, & Tsotsos, 2005). Like the term "attention", "selection" seems to have acquired multiple possible meanings:

> "Any kind of task-dependent modulation of sensory (or sensory-motor) neuronal responses? Spatial orienting? Selective facilitation? The gating out of unwanted information? Attentional tagging? Selective feature integration? Selective delay of processing? Making information "available to control responses and conscious experience"? "Entry to a limited-capacity short-term memory store"? All of these concepts, and many more, have been used as referents of the term *selection*." (Allport, 1993).

Allport (1993) argues that unless it can be shown that all these concepts refer to the same underlying mechanism, it is more likely that selection mechanisms act at several stages. In fact, this claim is supported by much experimental evidence from imaging and physiology studies (see section 1.4). Traditionally it was held that information processing occurs in a system limited in processing capacity (e.g. Broadbent, 1958). From this idea arises the need for selective attention, which allows only part of the present information to pass to more costly processing in the brain, such as matching visual input to object representations. Furthermore, this selection must come before the system's capacity becomes limited, i.e. it must come before some kind of "bottleneck" (Allport, 1993). In order to account for a fast selection process that is not yet limited in capacity, many models of selective visual attention include a preattentive stage of processing at which large parts of the visual input are processed in parallel. Preattentive is opposed to attentive processing, which involves shifts of focused attention (Treisman & Gelade, 1980). Most theories of visual search that followed the groundwork laid by Treisman and Gelade's (1980) feature integration theory (FIT) are, therefore, also general theories of selective attention (Müller & Krummenacher, 2006). The computational model of attention and item selection during visual search by Koch and Ullman (1985) adopts many ideas from FIT, and was later implemented in a neural network model by Itti, Koch, and Niebur (1998).

1.2 Models of attentional mechanisms during visual search

1.2.1 Feature integration theory, features, and search types

Feature integration theory (Treisman & Gelade 1980) proposes that there are two functionally independent sequential processing stages, an early preattentive and a later attentive stage, that are concerned with processes of visual perception. At the preattentive stage several “feature detectors” register in parallel the presence of a particular feature in the visual environment. A feature dimension, as defined by Treisman & Gelade (1980, pp. 98-99), is “the complete range of variation which is separately analyzed by some functionally independent perceptual subsystem”, whereas a feature is a particular value on that dimension. For example, a feature would be “red” if “color” is the perceptual dimension. “Vertical” would be a value on the “orientation” dimension. However, features do not uniquely correspond to physical properties of the visual input. These dimensions are perceptual, rather than purely physical, which implies that some higher-order properties, such as symmetry or homogeneity, may be analyzed by the perceptual subsystem. Consequently, it should be efficient to search for symmetry among non-symmetric patterns. More recent investigations have suggested density in the sense of a ratio of counts (Taylor & Badcock, 1988) and topological structure (Chen, 1982) as possible dimensions, which cannot be called physical properties per se. A comprehensive summary of dimensions ordered with respect to the likelihood of their status as features or basic attributes can be found in (Wolfe & Horowitz, 2004) and somewhat revised in (Wolfe, 2014). These authors concluded that there are between one and two dozen feature dimensions (Wolfe & Horowitz, 2004; Wolfe, 2014). Among the least disputed examples are color and orientation (Wolfe, 1998) as well as motion (Wolfe, 2014).

In the originally dichotomous view of search types (e.g. Treisman & Gelade, 1980) two paradigmatic versions of the task were distinguished: “feature search” and “conjunction search”. In feature searches the target is the item with a unique feature, such as a red square among green squares. Searching for a red triangle among red squares and green triangles is a conjunction search, since the target is defined by a unique conjunction of the feature dimensions it shares with each set of distractor items (here: a conjunction of color and shape). In most feature searches reaction time does not increase with increasing numbers of distractors. Plotting the time needed to find the target as a function of set size—the number of target and distractor items in the display—is an indicator of search efficiency. Searches in which reaction time does not depend on the number of distractors yield flat slopes. Thus, a flat slope indicates very efficient search, whereas a steep slope means that practically each item has to be checked one after another (in a serial manner), so this type of search is called inefficient. Most conjunction searches defined as above are inefficient. It is now generally accepted that there is no clear cut between these two types, but that, instead, some searches are somewhere in-between the efficient and inefficient classes. Thus, it has been suggested to view visual search as producing a continuum of search efficiencies (Duncan & Humphreys, 1989). Furthermore, this distinction is on a dimension apart from the search being a feature or conjunction search. Thus, there can be inefficient feature searches and efficient conjunction searches. Search efficiency has been shown to increase the more similar distractors are to each other and the more dissimilar they are from the target (Duncan & Humphreys, 1989).

FIT accounts for the finding of efficient feature search with preattentive processes that work in parallel over the entire visual field. Feature detectors feed the registered features of an object into separate “feature maps”, each of which codes only for one particular dimension. Initially, there is no coordination between different feature maps. The output of several feature maps “activates” locations on a master map, upon which attention sweeps like a spotlight in a serial manner from one location to another. Activations on the master map represent the presence of salient discontinuities present in the corresponding feature maps, without representing, however, what features of the visual input cause the discontinuities (Quinlan, 2003). In order to recover the features, a scan of focal attention along the master map is needed. Thus, after different features of an object are separately coded in different feature maps, they are “re-combined” at a later stage. Re-combination or integration of features requires the “glue” of focal attention. The notion of feature integration is central to the theory, as suggested by its name. In the original account spatial information of a given feature in the visual input is not directly accessible. Instead, observers could report the presence of features in the visual input but not their location. Correctly localizing features requires additional processing of the input under focal attention (Treisman, 1985). However, others suggested that not only could the preattentive stage promote detection of targets, but also their localization (Sagi & Julesz, 1985). Treisman (1988) later revised FIT to include spatial information in feature space.

1.2.2 Criticisms on the preattentive stage of processing

The distinction between a preattentive and an attentive processing stage has a long tradition (e.g. Neisser, 1967; Sagi & Julesz, 1985), and it has been useful in accounting for a variety of visual search findings. However, the two-stage architecture has also been criticized (Di Lollo, Kawahara, Zuvic, & Visser, 2001; Nakayama & Martini, 2011; Nakayama & Joseph, 1998; Wolfe, Treisman, & Horowitz, 2003). Some of the criticisms offered are:

1. The dualistic viewpoint presented in two-stage models of attention suggests a clear separability of search times into “parallel” and “serial” classes which conflicts with the continuum of search efficiency found across many visual search studies (D iLollo et al., 2001). Note that this criticism does not condemn the existence of a preattentive processing stage if its parallel character could be defined as a matter of degree.

2. The list of potential preattentive feature dimensions as revealed by parallel search has grown beyond biological plausibility (Di Lollo et al., 2001; Nakayama & Joseph, 1998). That is the case as long as it is assumed that parallel search is diagnostic of primitive (i.e. preattentive) feature dimensions and that there is a separate retinotopic feature map for

each dimension somewhere in the brain, both of which were promoted initially by FIT (Nakayama & Martini, 2011). As was briefly discussed above, identifying what counts as a primitive feature has been a difficult matter (cf. Wolfe & Horowitz, 2004). On the one hand, there are some examples of rather complex features that can evoke parallel search, such as digits (e.g. Wang, Cavanagh, & Green, 1994). Again, this criticism does not attack the existence of a psychological preattentive stage, unless a connection to biology is drawn. Indeed, the linkage between the properties of receptive fields in early visual cortex and the idea of feature analyzers at the preattentive stage has been made by many authors either implicitly (e.g. Treisman & Gelade, 1980) or explicitly (e.g. Sagi & Julesz, 1987; Nothdurft, 2005; Li, 2002). If this stage is thought of as part of a psychological model rather than as synonymous to lower level visual brain areas, we are left with the freedom to search for its biological basis in the whole brain.

3. It has been shown that even otherwise easy search tasks involving only primitive features such as line-orientations can yield steep search slopes when attention is divided between two simultaneous tasks (Joseph, Chun, & Nakayama, 1997). Joseph et al. (1997) showed that performance in a feature search task was slowed when observers had to simultaneously report a cued letter in a rapid serial visual presentation letter stream while also doing the visual search task. The drop in performance, compared to doing only one of the tasks, was interpreted as due to shared attentional resources between the two tasks in the simultaneous condition. Thus, the “preattentive” processes seem to depend on attention after all, which argues against its original conceptualization and “bring[s] into question the necessity, or even the usefulness, of postulating an encapsulated preattentive stage” (Di Lollo et al., 2001, p.480). In contrast, Wolfe (2003) even argues for a tautological assumption of a preattentive stage, as long as attention is imagined as a selective process, with some processing preceding the selective stage.

To summarize, none of the criticisms sketched above argues against the existence of a preattentive processing stage per se, but against some specifications of the concept. Preattentive or preselective (Wolfe, 2003) processing is the most common explanation for bottom-up control over attention and efficient visual searches. Where this stage is implemented in the brain is not clear; lower level vision is a suggestive but not a mandatory hypothesis.

1.2.3 Models after FIT

Several theoretical accounts rely on the concept of a saliency map and on that of a competition between neurons to explain the mechanism of bottom-up attention. The saliency map is a proposed mechanism that accounts for shifts of visual attention (Itti & Koch, 2000). It is an explicit two-dimensional map that encodes stimulus conspicuity, or saliency, at every location in the visual scene (Itti & Koch, 2001). It is topographically organized, with every location on the map corresponding to a spatial location in the visual input. Activities on the saliency map represent visual saliency topographically, irrespective of the feature dimension that makes a particular location salient. Thus, saliency map activations do not reflect whether a feature is salient for its unique color, orientation etc. in the image. Attention is then drawn to the most salient locations (although most models do not specify how this happens or what attention really is). In some models, the saliency map depends solely on “bottom-up” factors (e.g., Li, 2002; Koch & Ullman, 1985). This is the most wide-spread definition of saliency. Some authors prefer the term “priority” in order to refer to the combined bottom-up saliency value and behavioral relevance of a stimulus (Serences & Yantis, 2006; Fecteau & Munoz, 2006). The following section will present some examples of models of selective attention and visual search; two of the first kind, and one that models how attentional selection is guided by both bottom-up and top-down factors (i.e., “Guided Search” model, Wolfe, 1994).

1.2.3.1 Koch and Ullman’s model

One such model is the theoretical neuron-like network from Koch and Ullman (1985) that aims to be compatible with physiology and anatomy. On the one hand, it was motivated by the psychological two-stage models (e.g. Neisser 1967; Treisman & Gelade 1980; Julesz, 1984; Ullman, 1984), in which a preattentive mode (or stage) processes features in parallel, and an attentive mode (stage) is required for recognition of objects. On the other hand, it was inspired by the argument that for computational reasons the analysis of visual information must become serial beyond a certain stage.

Individual feature maps code for a particular feature dimension each, such as color or orientation, inspired by early physiological findings (e.g. Zeki, 1978). For example, primary visual cortex and adjacent visual areas all show topographical organization but differ with respect to the distribution of orientation- and color-selective cells amongst them. The maps are not necessarily assumed to be in physically different locations (Koch & Ullman, 1985). The conspicuity or saliency of a location in the visual scene determines how features that are present at this location will be represented in the topographic feature maps. More precisely, the more salient a location, the higher the activity it evokes in the corresponding location on the different feature maps. The authors are reasonably vague about a possible place of the saliency computation in the brain, but they suggest a possible physiological mechanism: “Local, inhibitory connections, mediating lateral inhibition, occur either at an earlier stage or within the feature maps. Thus, locations that differ significantly from their surrounding locations are singled out at this level.” (p. 220).

Several mappings happen from the stage of the visual input to the final stage in the model. First, visual input features are mapped onto feature maps. Then the outputs of different feature maps are combined in a saliency map, which represents a global measure of saliency of locations in the visual field. According to a so-called “winner-takes-all” operation, all activations on this map are set to zero except for the highest one in order to route the properties of a single selected location onto a more “central non-topographic representation”, which denotes some stage of cortical processing further removed from the periphery (p. 220). This mapping from early representations on feature maps onto the central representation is referred to as the “expression of early selective visual attention” (p. 219). Similar to FIT, it is suggested that attention “fuses” the information from different maps into a “coherent whole” (p. 219).

1.2.3.2 V1 as a “bottom-up saliency map”

Unlike most other models, the saliency map in this model is not removed from the stage of feature registration. The idea is that primary visual cortex (V1), due to the known properties of its cells, is capable of signaling salience of a subset of visual input features, without the need of separate feature maps or a subsequent combination of them (Li, 2002). V1 cells show preferences in their firing behavior toward specific edge-orientations, wavelengths, and the eye-of-origin of the visual input (left/right eye stimulation). In this computational model, saliency of a visual location describes the location’s ability to attract attention without top-down factors (Zhaoping, 2005). The saliency value is computed regardless of the feature presented at a given location and is quantified in terms of neural firing rates. For example, the saliency of a red dot may be compared with that of a vertical moving bar, if these two features evoke the same neural firing rate. The saliency of a given location depends both on the visual input strength (or contrast) at that location, and its near context outside the classical receptive field (CRF) of cells at this location. The CRF of a particular neuron in visual cortex is the region of visual space in which stimulation will evoke a response in this neuron. Responses can be evoked from several cells with overlapping CRFs at one location, some of them tuned to the input color, some tuned to its orientation, etc.

The notion that saliency computations are possible in V1 has also been considered by Nothdurft (2005). Salience, in this account, is “an immediate perceptual impression” that depends on local feature contrast, which is a general term to describe how a target item differs from its neighboring items along a dimension (e.g. luminance, color, spatial frequency, motion, depth) (Nothdurft, 2005). If sufficiently different from its neighbors, the target will stand out. However, if neighboring items themselves vary in their properties, the salience of the distinct target will be reduced. Nothdurft proposes that salience from feature contrast is reflected in the population response of V1 neurons, based on the fact that the population response to the target varies in qualitatively the same way as the observer's subjective perceptual impression of target salience within different surrounds. V1 neurons respond differently when presented with or without feature contrast—this is called contextual modulation—possibly due to lateral interactions between neurons within the same area, or to feedback from subsequent areas (Nothdurft, 2005). Li (2002) is explicit about the kind of lateral interactions between groups of V1 neurons (e.g. cortical hyper columns) that could produce a saliency map and shows how a computational model of saliency computations in V1 explains visual search behavior.

1.3 A model that includes top-down control

In the models described above saliency is defined to depend only on stimulus features. There are other models that propose a mechanism for other factors to influence or bias attentional selection, collectively called top-down factors.

1.3.1 Guided Search

One such model is the Guided Search (GS) model (Wolfe, 1994; Wolfe, Cave, & Franzel, 1989; Wolfe, 2014). GS is very similar to FIT. The main addition to the FIT model is that attention could be guided by prior knowledge of the target properties (Wolfe, 2014). Consequently, even serial or inefficient search can be made more efficient via top-down guidance, as compared to what would be predicted from the strictly dichotomous FIT architecture. What essentially guides attention are a set of basic attributes extracted from the early stages of visual processing (Wolfe, 2014). For each attribute, two forms of guidance are possible: bottom-up and top-down guidance. The former happens when the difference between a visual item and its neighboring items attracts attention. This is essentially the same as some formulations of salience (e.g. Nothdurft, 2000). Attention is guided by the latter type when the observer is told to look for objects having a specific preattentive feature (Wolfe, 2005). Contributions from each feature are combined in an activation map. An expectation for the target features will preactivate corresponding feature maps in a top-down fashion before the outputs of feature maps are combined in the activation map (Wolfe, 1996).

There may be many types of top-down control over attention during visual search apart from guiding attention to the instructed target. Two further examples are illustrated below.

1.3.2 Top-down allocation of attention due to observer experience

Recently experienced stimuli can influence subsequent allocation of attention. If a novel target shares features with preceding search targets (even response-irrelevant features), it is found more effectively (Maljkovic & Nakayama, 1994; 2000). Priming effects can be considered an example of automatic top-down control, since something about the history of the observer has changed, rather than the stimulus itself (Wolfe, Horowitz, Kemner, Hyle, & Vasan, 2004). Another example of automatic processing that depends on the observer's history is contextual cueing. Contextual cueing in visual search tasks refers to the situation where reaction time of target detection is faster for displays with repeated spatial configurations of target and distractors compared to displays with novel configurations. The process responsible for guiding attention quickly to the target within repeated contexts may be seen as top-down knowledge from experience (Chun & Jiang, 1999).

1.3.3 Top-down processes that impede efficient search

Smilek, Enns, Eastwood & Merikle (2006) were motivated by the fact that most visual search studies neglect the influence of "cognitive strategy", a process that is under the observer's voluntary control but is independent of her knowledge of the task. By giving observers (a) the instruction to give up active control over guidance of attention, or (b) a concurrent memory task, their search performance became more efficient on a hard search task (for a circle with a small gap among circles with two small gaps) but not in the easy version (circle with a large gap among circles with two small gaps). This finding allows for two conclusions: Sometimes "too much" top-down processing can be detrimental to a search task. Importantly, this top-down processing is to some degree subject to the voluntary control of the observer.

1.4 The locus of selection in the brain

Saliency computation is the mechanism by which selective attention is supposed to choose its target for processing. As discussed in the preceding sections, salient features may include values such as color or orientation embedded in a particular context, but also behavioral relevance, according to some definitions of saliency (e.g., Serences & Yantis, 2006). There are several brain areas that could mediate saliency computation and selective attention based on their involvement in visual search and other selective attention tasks. Inactivation of monkey superior colliculus leads to deficits in target selection (McPeek & Keller, 2004). Brain imaging studies show that posterior parietal cortex activates in tasks where attention is directed to one of multiple stimuli without prior knowledge of the target location (Corbetta, Shulman, Miezin, & Petersen, 1995). The frontal eye field (FEF) might form a map of salience according to behavioral relevance of a stimulus (Bichot & Schall, 2005), something which has also been suggested to occur in lateral intraparietal cortex (Colby & Goldberg, 1999; Bisley & Goldberg, 2010). The fact that FEF responses were selective to whether a stimulus was relevant to a visual search task (the target), conspicuous by intrinsic properties while not being the target (singleton distractor), or a non-salient distractor, lend support to the hypothesis that FEF neurons signal behavioral significance rather than the saccade goal only. The human homolog of FEF is involved in both shifts of attention that are accompanied by eye-movements and those that are not (Corbetta & Shulman, 2002). Other areas involved in visual search are area MT (Treue & Martinez-Trujillo, 2006), the pulvinar (Shipp 2004), and dorsolateral prefrontal cortex (Nobre, Coull, Walsh, & Frith, 2003). Moreover, in this latter study parietal cortex was particularly activated during inefficient searches, independently of whether the target was defined by a unique feature or conjunction of features. This suggests that parietal cortex contributions to inefficient visual search involve other functions apart from feature binding.

Two recent studies showed that target relevance alters neurons’ responses in area V4, suggesting that this ventral stream area might combine bottom-up and top-down saliency signals. Bichot, Rossi, and Desimone (2005) found that neurons in monkey V4 showed enhanced responses for objects that shared a feature (color, shape, or both) with the target in a complex visual search array. Furthermore, Ogawa and Komatsu (2004) showed that feature selectivity and task-related modulations occur in V4 cells of monkeys during a version of the additional singleton search task (cf. Theeuwes, 1991). Based on the known properties of V4 cells, i.e. their shape- and color selectivity, and the fact that their responses can be influenced by attention to features of visual objects, the authors investigated how visual signals encoding feature contrast between objects are modulated by attention specific to a particular dimension (color or shape). The stimuli consisted of 6 items: 2 singletons, one unique in color, the other unique in shape, and 4 nonsingleton (non-unique) items. In each trial, the singleton of the ongoing instructed dimension was the target. The other singleton was referred to as distractor in that trial. Displays were arranged so that the receptive field of V4 cells contained a target, distractor, or nonsingleton in a given trial. Monkeys were required to make a saccade toward the target. Population responses of 36 neurons were higher when the item in their RF was a singleton in the current searched-for dimension (target), compared to when it was not (distractor and nonsingleton, respectively). This finding shows that V4 responses carry information about the behavioral context. The cells were furthermore selective for shape and color. Together, these findings suggest that sensory and behavioral information are combined in V4 neurons' responses.

Primary visual cortex is not among the areas commonly associated with saliency computation, despite the model which shows how V1 could create a bottom-up saliency map (Li, 2002). Indications of this proposal from neuroimaging are, as yet, sparse (e.g., Zhang, Zhaoping, Zhou, & Fang, 2012; Melloni, van Leeuwen, Alink, & Müller, 2012). Finally, the involvement of one or more distinct areas in saliency computations does not preclude a role for other areas in saliency computation:

> "[T]he idea of a unique, centralized saliency map appears today to be challenged by the multiplicity of candidate neural correlates recently unraveled, including areas in the lateral intrapariental sulcus of the posterior parietal cortex, the frontal eye fields, the inferior and lateral subdivisions of the pulvinar, and the superior colliculus (Kustov & Robinson, 1996; Gottlieb et al., 1998). One possible explanation for this multiplicity could be that some of the neurons in all of those areas indeed are concerned with the explicit computation of salience, but are found at different stages along the sensory-motor processing stream." (Itti, 2005, p. 579).

1.5 Studying attention through eye-movements

Eye movements are a manifestation of the allocation of attention over visual input. Eye-tracking is a means that can be used to study spatial distributional and temporal aspects of visual attention by recording movements of the eye. Eye-tracking devices usually have a sampling frequency between 25 and 20000 Hz and with spatial accuracy usually around 0.5° visual angle (Andersson, Nyström, & Holmqvist, 2010). It is non-invasive, so that subjects are not impeded or influenced in their movement or perception.

For most eye-tracking studies, two categories of eye movements are particularly relevant: saccades and fixations. Through saccades a visual object is placed under the fovea, where visual acuity is highest. During saccades, one is practically “blind” to any changes in the visual scene (McConkie & Currie, 1996). Visual stimuli can be perceived only during fixations, and there is a relationship between fixation duration and properties of the stimulus: fixations are affected by luminance and contrast, but also by how informative the fixated region or object is (Henderson, 2003). Therefore, the effects of cognitive or semantic properties and those of visual properties of a stimulus on fixation behavior may look similar to the experimenter. On the other hand, a fixation does not automatically mean conscious processing of the underlying object. That is especially the case with very brief fixations (due to brief stimulus exposure times followed by a mask, or because the observer’s attention might have been drawn to the object and subsequently diverted), before the information about its identity could enter conscious processing. Apart from saccadic eye movements, other ocular parameters can be used, such as pupil dilation to gain information about the observer’s arousal state (e.g. Kahneman & Beatty, 1966).

Often, shifts of attention and eye movements occur in tandem; however, they are not the same. Eye-movements can be an expression of overt orienting—selectively attending to a stimulus by moving the eyes to point into that direction (Posner, 1980). In addition, attention can be deployed in a covert fashion, meaning, without moving the eyes. For example, in the precueing paradigm, attention is first paid to a spatial location well in advance before the subject saccades toward it (Posner, 1980). In order to study covert attention shifts in human subjects, however, eye-tracking does not suffice as a method. In particular, experimenters sometimes desire to study covert attention without "contamination by eye movements" (e.g. Folk, Remington, & Johnston, 1992). The latter are prevented by very short stimulus presentation times which do not allow for saccades to occur. However, unless observers are explicitly instructed to prevent eye movements, they show a natural preference to move their eyes in search tasks (Findlay & Gilchrist, 1998). But do all gaze shifts imply a shift of attention between spatial locations or objects? It is generally acknowledged that covert shifts of attention obligatorily precede saccadic eye movements (Hoffman & Subramaniam, 1995; Kowler, 1995; Deubel & Schneider, 1996). While this research focused on voluntary saccades, there is also behavioral evidence that covert shifts of attention precede involuntary saccades (Peterson, Kramer, & Irwin, 2004). Interestingly, neuroimaging findings suggest substantial overlap between neural systems that mediate attentional shifts and those underlying saccadic eye movements (e.g. Beauchamp, Petit, Ellmore, Ingeholm, & Haxby, 2001; Nobre, Gitelman, Dias, & Mesulam, 2000). Thus, what attracts the eye can be said to have attracted the observer's attention. The pattern of saccades revealed by eye-tracking gives information about which elements of the visual stimulus or environment attract the observer's attention, in what order, and how often (Yarbus, 1967). It is then another task for the experimenter to interpret this pattern, as there can be many different reasons why a visual object draws attention to it.

Chapter 2 Literature review: Visual search in autism

2.1 Chapter overview

This chapter reviews studies of visual search in people with autism. The following two sections in this chapter are meant as an introduction to research on visual processing in autism, starting with a definition of autism spectrum disorder. Section 2.4 will discuss the theoretical groundwork related to visual processing in autism. In particular, it is discussed how the outlined theories relate to visual search in autism, and which of the constructs put forward by them are of relevance in this context. This is followed by a detailed review of visual search findings in autism in sections 2.5 and 2.6.

2.2 Definition of autism spectrum disorders (ASDs)

Autism spectrum disorder (ASD) is a life-long developmental disorder characterized by deficits in social interaction and communication abilities as well as by restricted interests and repetitive patterns of behavior (APA, 2013). It is currently estimated that 60 per 10000 individuals in developed countries are diagnosed with ASD, with a higher proportion of males being affected compared to females (Newschaffer et al., 2007). The Diagnostic and Statistical Manual of Mental Disorders (DSM-5; APA, 2013) includes within ASDs the previously separate categories of autistic disorder (autism), Asperger's disorder, childhood disintegrative disorder, and pervasive developmental disorder not otherwise specified (PDD-NOS). The etiology of ASD, which has a strong genetic basis and possibly also involves environmental factors, remains to be fully understood (Abrahams & Geschwind, 2008). In this thesis the term autism will be used interchangeably with ASD.

2.3 Vision in ASD

Within the last fifteen years interest has grown in sensory processing abnormalities in autism. In particular, a large body of research in visual processing has been dominated by the findings of abnormal processing of faces and face-related stimuli (review by Behrmann et al., 2006), impaired coherent motion perception (Spencer et al., 2000;

Milne et al., 2002; Pellicano, Gibson, Maybery, Durkin, & Badcock, 2005), and differences in several aspects involving attention (review by Ames & Fletcher-Watson, 2010), as well as the superior behavioral performance in visuospatial tasks in autism (some reviewed by Dakin & Frith, 2005). However, many findings have been contradictory and remain a topic of further research, such as findings in detection thresholds of coherent dot motion (e.g. Spencer et al., 2000; but see Jones et al., 2011), susceptibility to visual illusions (Happé 1996; Mitchell, Mottron, Soulières, & Ropar, 2010; but see Ropar & Mitchell 1999; 2001), performance on hierarchical figures tests, the spatial spread of attention and visual orienting (Simmons et al., 2009). Raised motion coherence thresholds in ASD are suggestive of a dorsal stream deficit for some (Spencer et al., 2000). More recently, some studies investigated form processing in the ventral stream with non-social stimuli in ASD (Milne et al., 2006; Vandenbroucke et al., 2008), and few studies find difficulties with object perception in autism (briefly reviewed by Behrmann et al., 2006). Based on findings regarding both pathways, it has been suggested that visual processing in ASD might be connected with difficulties to integrate information at the higher level of both visual pathways, with less available evidence with regard to the ventral pathway (Grinter, Maybery, & Badcock, 2010). Orthoptic measures show mostly unaffected binocular visual function in autism (Milne, Griffiths, Buckley, & Scope, 2009), although two studies report a high prevalence—21% and 50%, respectively—for strabism (Scharre & Creedon, 1992; Kaplan et al., 1999). Findings of eye-movement measures are few and inconsistent (for a review, see Brenner, Turner, & Müller, 2007). There is no evidence for problems in contour-integration or depth perception in people with ASD (Simmons et al., 2009). Findings of higher visual acuity in autism can be considered invalidated (Bach & Dakin, 2009; Falkmer et al., 2011). The bulk of evidence shows normal static contrast sensitivity in autism compared with typical observers (Simmons et al., 2009).

2.4 How cognitive theories of autism relate to visual search

Superior performance in difficult visual search of people with autism compared with typical control groups is a highly replicated finding in children (e.g. Plaisted, O'Riordan, & Baron-Cohen, 1998; O’Riordan, 2000; O’Riordan, Plaisted, Driver, & Baron-Cohen

2001; Kaldy, Kraper, Carter, & Blaser, 2011; but see Keehn, Shih, Brenner, Townsend, & Müller, 2013) and in some studies with adults with autism (e.g. O'Riordan, 2004; but see Constable, Solomon, Gaigg, & Bowler, 2010). Most studies use reaction time (RT) for target detection as measure of performance. Three influential theories of autism address findings of visual search and other visuospatial tasks. First, the enhanced perceptual funcionting (EPF) model (Mottron & Burack, 2001; Mottron, Dawson, Soulières, Hubert, & Burack, 2006) and its extension, the Veridical Mapping model (Mottron et al., 2013) hold that "overdeveloped" low-level perception and increased autonomy of lower level brain areas from top-down processes account for the performance of people with autism in visual (and other) tasks. People with ASD are predicted to excel on visual search, by this account, due to enhanced pattern discrimination and –matching mechanisms that operate in low- and mid-level vision (Mottron et al., 2013). Another theory is the reduced generalization theory (RGT), which accounts for superior performance in some visual tasks with an inability of people with autism to recognize the similarities between stimuli or situations (Plaisted, 2001). According to this theory, properties that are held in common between visual objects are processed relatively poorly in autism, whereas properties unique to an object (i.e., discriminating features) are processed relatively well. This leads to reduced generalization and enhanced discrimination abilities at perceptual, attentional and, only consequently, at conceptual levels. EPF and RGT both account for superior visual search in autism compared to typical people by proposing a somehow enhanced ability to discriminate the target from distractors. Third, the weak central coherence (WCC) account, suggests that people with autism have difficulties integrating information into a gestalt or "whole" (Frith, 1989) or will at least tend not to do so by default, unless explicitly instructed to (Happé & Frith, 2006). According to the WCC account, whereas the normal cognitive system strives for "central coherence" over a wide range of stimuli and contexts, this drive is weak in ASD (Frith, 1989). This account predicts difficulties for people with autism in tasks requiring "central" processing compared to "peripheral" or sensory processing, regardless of the sensory modality involved (vision, audition, etc). Similar to EPF, WCC also claims that superior performance will be found in tasks with a disadvantageous "capture" by gestalt and/or where focus on detail is of advantage, such as the Embedded Figures Test (EFT) or Block Design (BD) tasks (Happé & Booth, 2008). The EFT requires spotting of simple geometric shapes embedded within more complex line drawings. In the BD task, arbitrary patterns have to be constructed from simple elements, by breaking down the larger pattern into its building blocks. The implications of the WCC account for visual search performance of people with ASD are not straightforward; at least attention to detail, which is predicted for autism by WCC, should be advantageous for conjunction search tasks. On the other hand, one might think that WCC predicts conjunction search to be inferior in autism, as this task requires integration of features of an object (e.g. color and form). In contrast, conjunction search in autism is often found to be equal or superior in performance to typical people (see section 2.5 below). However, this is not necessarily inconsistent with WCC theory, as the type of integration or binding of features to conjunctive search targets is not necessarily the type of integration of parts into wholes addressed by the original WCC account (further discussed in section 3.2.1).

2.5 Studies of visual search in ASD

Performance of individuals with autism in visual search tasks is often described as “enhanced” or “superior” with levity. Differences among the many kinds of visual search tasks and their implications are thereby disregarded. To the best of my knowledge, only one article dedicated to the review of studies of visual search in autism has been published so far (Kaldy, Giserman, Carter, & Blaser, 2013). Their self-declared “mini-review”, comprising 22 studies published in the last 15 years, concludes that, overwhelmingly, individuals with ASD perform better, as measured by lower RTs in most studies, compared with controls, and that this performance advantage is present across ages, the spectrum of severity, and task designs. As Kaldy et al. do not disclose their inclusion criteria for the studies reviewed, it is unclear why at least three recent studies of visual search in ASD are missing (i.e., Kemner, Van Ewijk, Van Engeland, & Hooge, 2008; Constable et al., 2010; and Keehn et al., 2013). Another aspect of their review is that results for feature search tasks are systematically ignored. For the purpose of this chapter we will refer to those tasks as "feature search" where the target can be identified by its difference from distractors on a single perceptual dimension according to Treisman's examples (Treisman & Gelade, 1980). "Conjunctive" or "conjunction" search means that the target can only be identified by its unique combination of features it shares with subsets of distractors. If a distinction is made between feature search and conjunctive search tasks, the answer to the question of whether people with ASD are generally faster in visual search becomes more complex.

Exclusion criteria of the present review were:

- No studies using "social" stimuli (e.g. faces, face drawings, photographs of people, etc.) were included
- Studies diverging too much from classic visual searches, due to their research objective, were also excluded (such as preview-search, as in Keehn & Joseph, 2008)
- Only those studies were included with a clinical group, excluding studies where the participant groups were formed solely by typical observers grouped by their autism spectrum quotient (AQ) score

The AQ is a measure of autistic-like traits in the normal population (Baron-Cohen, Wheelwright, Skinner, Martin, & Clubley, 2001). It should be mentioned, however, that some studies with adults of the normal population yield faster visual search in high versus low AQ scorers (e.g. the last four studies listed in table 1 of Kaldy et al., 2013). Note that this is not a systematic review, however. Not including studies with "social" stimuli raises an interesting question: Do visual search results differ between studies employing social stimuli and those with non-social stimuli? This question must be answered by future research. Here, the objective is to consider visual search performance free of any possible confounding factors, such as atypical processing of faces in autism. Moreover, by including studies that were not present in the mini-review, I am going to challenge the conclusion that the performance advantage is present across ages. In particular, with regard to adults visual search is not clearly superior in ASD.

The contribution of this review is to provide novel conclusions obtained by sorting the findings according to search designs (feature or conjunction) and stimuli employed. Note that O’Riordan, Plaisted, Driver, and Baron-Cohen (2001) already concluded that conjunction search is superior in autism, and feature search only if it is difficult enough so that ceiling effects in performance are avoided. Furthermore, they note that differences between autism and control groups are sometimes confined to target absent trials, but not always. Importantly, however, their conclusions were based on few available studies at that time (e.g. Plaisted et al., 1998; O’Riordan et al., 2001). The more recent studies have to be judged with respect to this particular pattern of results. This is particularly so for studies on adults with autism, which are still rare but more numerous than in 2001. Moreover, neither O’Riordan et al. (2001) nor Kaldy et al. (2013) commented on the types of search stimuli typically employed.

Table 1 explores the question of whether visual search yields faster RTs in ASD with respect to feature and conjunction searches. The “Search type” column shows whether a feature or conjunction search was employed (if both were used in the same experiment their respective results are listed in separate lines of the table). The “Result” column indicates whether RTs in the autism group were faster compared to typical, i.e., if the performance of the ASD group was higher compared with the performance of the TD group ("ASD > TD"), or if they were not ("ASD = TD"). Notably, none of the studies found significantly worse performance in the ASD group compared to typical. Studies are sorted from top to bottom by chronological age of their clinical participants. The stimuli employed in these studies were classified into five categories (A-E; see explanation at the bottom of Table 1). There is a general trend of faster RTs by ASD compared to TD in conjunction search. This is particularly so the more difficult the task, i.e. groups differ more with increasing set sizes and in target-absent trials, and the more the search falls into the category of inefficient searches. However, with respect to feature search, studies do not overwhelmingly show an advantage for ASD compared to controls. The pattern of "ASD > TD" in conjunctive search and "ASD = TD" in feature search is particularly obvious (a) for participant groups of children below the age of 13, and (b) for versions of the "color-shape conjunction task", such as the search for a red letter X among red T and green X distractors (used e.g. by Plaisted et al., 1998). An exception to (a) is O'Riordan's (2004, Exp. 3) finding of faster conjunctive search in adults with autism. In fact, conjunctive search tasks have rarely been applied to adults with autism (of the studies listed here, there is only one more: Caron, Mottron, Berthiaume, & Dawson, 2006). Some studies with children or young adults used stimuli other than (b), such as items from groups of dots (Iarocci, Burack, Shore, Mottron, & Enns, 2006), Gabor patches (Baldassi et al., 2009), or Block-Design inspired squares (Caron et al., 2006, Exp. 4), and find "ASD = TD", consistent with the general trend in feature search. Interestingly, some difficult feature searches yield faster RTs for people with ASD, such as the search for a vertical line among tilted lines (O'Riordan, Plaisted, Driver, & Baron-Cohen, 2001; Collignon et al., 2013; Kemner et al., 2008; but see Hessels, Hooge, Snijders, & Kemner, 2014).

The lack of performance advantage of people with ASD in feature search (especially in children) might be partly due to "floor effects" in RT, i.e. because searches are so easy, RTs are so low in both groups that there is no room for differences. This leads to the prediction that, if the feature search is difficult enough, children with ASD should be faster than typical children. This hypothesis receives support from a finding of faster RTs in children with ASD in a difficult feature search for orientation (O'Riordan et al., 2001, Exp. 2).

One caveat of the frequently employed color-shape conjunction search is that performance advantages might reflect a more successful strategy, rather than a perceptual difference between participant groups. Most of the studies in children employed a version of this task (see Table 1, stimulus key A). In most of them, the target remained the same over trials, e.g. always a red letter X. It is known that subjects can adopt a strategy of reducing search to subsets of distractors in conjunction search, such as the subset of red distractors or the subset of X-shapes in this case (Egeth, Virzi, & Garbart, 1984). The faster reaction times of the autism group might, therefore, reflect a better task strategy, rather than a difference in the perceptual capabilities between the groups. However, such a strategy would not be helpful in a situation where the target identity is varied across trials, e.g. either a red X or a green T (distractors are always red Ts and green Xs). Children with autism demonstrated faster RTs than controls even in this situation (Exp. 2, O’Riordan, 2000).

Table 1

Summary of findings of visual search in autism spectrum disorder

Reference	Search type (feature/ conjunction)	Result (ASD = TD / ASD < TD / ASD > TD)	Stimulus key	Stimulus details	Participant ages in years (ASD;TD)*	Sample size (ASD; TD)
Kaldy et al., 2011	Feature	ASD = TD	A	red/blue apples, elongated shapes	4.8; 2.5	17;17
	Conjunction	ASD > TD	A			
Iarocci et al., 2006,				Exp.1: search items are groups of four black and/or white displaced or aligned dots; Exp. 2: search items are hierarchical figures	7.9; 9.3, 8.0 (Exp.1) 7.8; 8.4, 7.8 (Exp.2)	12; 12 (Exp.1) 20; 12 (Exp.2)
Exp.1	Feature	ASD = TD	B			
Exp.2	Feature	ASD = TD	B			
O'Riordan et al., 2001				Exp.1: red S target among red T and green X; Exp. 2: 18° tilted line among vertical lines (easy) or vertical line among 18° tilted lines (hard)	8.4; 8.2 (Exp.1) 8.8; 8.3 (Exp.2)	12; 12 (Exp.1) 12; 12 (Exp. 2)
Exp.1	Feature	ASD = TD	A			
	Conjunction	ASD > TD	A			
Exp.2						
easy	Feature	ASD = TD	C			
hard	Feature	ASD > TD	C			
O'Riordan, 2000				Exp. 1 & 2: letters, varying colors and identities	9.2; 8.6 (Exp.1) 9.4; 8.5 (Exp.2)	11; 12 (Exp.1) 12; 12 (Exp.2)
Exp.1	Conjunction	ASD > TD	A			
Exp.2	Conjunction	ASD > TD	A			

Reference	Search type (feature/ conjunction)	Result (ASD = TD / ASD < TD / ASD > TD)	Stimulus key	Stimulus details	Participant ages in years (ASD;TD)*	Sample size (ASD; TD)
Plaisted et al., 1998	Feature Conjunction	ASD = TD ASD > TD	A A	Feature search for red S among red T and green X; conjunctive search for red X among red T and green X	8.8; 7.8	8; 8
O'Riordan & Plaisted, 2001				Exp. 1: red/green filled rectangles of different widths and orientations (task 1: color-orientation conjunction, task 2: color feature; task 3: color-shape-orientation conjunction) ; Exp. 2: letters, varying colors and shapes (all color-shape conjunctions; task 1 has lowest target-distractor similarity)	9.2; 8.6 (Exp.1) 9.0; 8.5 (Exp.2)	15; 15 (Exp.1) 13; 13 (Exp.2)
Exp. 1						
Task 1	Conjunction	ASD > TD	A			
Task 2	Feature	ASD = TD	A			
Task 3	Conjunction	ASD > TD	A			
Exp. 2						
Task 1	Conjunction	ASD = TD	A			
Task 2	Conjunction	ASD > TD	A			
Task 3	Conjunction	ASD > TD	A			
Task 4	Conjunction	ASD > TD	A			
Baldassi et al., 2009	Feature	ASD = TD	B	Gabor patches	11.2; 12.4	12; 15
Jarrold et al., 2005	Feature Conjunction	ASD > TD ASD > TD	A A	red/green clowns	12.4; 6.5	18; 18

Reference	Search type (feature/ conjunction)	Result (ASD = TD / ASD < TD / ASD > TD)	Stimulus key	Stimulus details	Participant ages in years (ASD;TD)*	Sample size (ASD; TD)
Keehn et al., 2013	Feature	ASD = TD	D	target T among rotations of itself	13.8; 14.0	19; 19
Joseph et al., 2009	Feature	ASD > TD	D	target L among Ts	14.6; 14.2	21; 21
Caron et al., 2006				Exp. 4: squares inspired by Block Design task	23.3, 18.9; 18.6, 16.9	8, 8; 8, 10
Exp.4	Feature	ASD = TD	B			
	Conjunction	ASD = TD	B			
Hessels et al., 2014				Exp. 1: vertical target line among 10° lines; Exp. 2: target was 17° tilted line among vertical lines (easy), or vertical line among 17° tilted lines (hard)	21.6; 21.6 (Exp.1) 22.9; 22.8 (Exp.2)	19; 31 (Exp.1) 13; 14 (Exp.2)
Exp.1	Feature	ASD = TD	C			
Exp.2						
easy	Feature	ASD > TD	C			
hard	Feature	ASD = TD	C			
Kemner et al., 2007				vertical/titled lines, same as O'Riordan et al. (2001)	22.1; 21.2	7; 8
easy	Feature	ASD > TD	C			
hard	Feature	ASD > TD	C			

Reference	Search type (feature/ conjunction)	Result (ASD = TD / ASD < TD / ASD > TD)	Stimulus key	Stimulus details	Participant ages in years (ASD;TD)*	Sample size (ASD; TD)
O'Riordan, 2004				Exp. 1: feature search for letter N among Q, P; "conjunctive" search for R among Q, P; Exp. 2: ellipse among circles; Exp. 3: letters, varying color and identity	Exp.1-3: 22; 22	10; 10 (Exp.1-3)
Exp.1	Feature	ASD = TD	D			
	Conjunction	ASD > TD	D			
Exp.2	Feature	ASD > TD	E			
Exp.3	Conjunction	ASD > TD	A			
Remington et al., 2009	Feature	ASD = TD	D	letters, arranged in circle with additional peripheral distractor	23.5; 26.6	14; 23
Remington et al., 2012	Feature	ASD = TD	D	same as Remington et al. (2009), with simultaneous detection task	24.1; 25.8	14; 14

Reference	Search type (feature/ conjunction)	Result (ASD = TD / ASD < TD / ASD > TD)	Stimulus key	Stimulus details	Participant ages in years (ASD;TD)*	Sample size (ASD; TD)
Collignon et al., 2013 *"tone absent" condition*	Feature	ASD > TD	C	red and green lines; target horizontal or vertical, distractors oblique; color of subset of lines switched at random	24.5; 21.0	16; 19
Constable et al., 2010	Feature	ASD = TD	E	ellipse among circles	37.6; 44.6	16; 16

ASD autism spectrum disorder, *TD* typically developing/developed control group, *Exp.* experiment. Stimulus keys: *A* version of a (classic) color-shape conjunctive or feature search, *B* rare stimulus, *C* line orientation feature search, *D* unicolor letter search, *E* ellipse among circles, *months were rounded to one decimal behind period. Ages of subgroups within ASD or TD groups are separated by commas.

2.6 Conclusion

Is visual search "superior" in autism compared to typical people? By examining Table 1, it becomes clear that the answer is more complex than is usually admitted. Children with ASD robustly outperform controls in conjunction but not feature search. This pattern of results has been noted previously (O'Riordan et al., 2001), and could be confirmed by the present review which includes more recent studies. Findings for adults are mixed with respect to feature search, and are largely lacking with respect to conjunction search. Differences between clinical and typical groups are more reliably revealed the more difficult the search. Future investigations should direct their attention to conjunction search in adults, and to difficult feature search designs in children, to further examine whether there are similar trends across ages.

Chapter 3 Experiment: Testing visual feature search under interference of top-down shape recognition in people with autism spectrum disorder

Mara Tribull, Sarah White and Li Zhaoping

3.1 Abstract

It has been suggested that visual attention in individuals with autism differs from typical vision. However, it is not quite clear which visual processes are affected. We compared high-functioning adults with autism spectrum disorder (ASD) with typical individuals on a fast feature detection process in a visual search task and a slower top-down process that interferes with the detection task. Observers searched for the target shape containing a uniquely oriented bar. The target shape was embedded within non-target shapes, but to solve the task, shape information can be completely ignored. However, the fact that non-targets were reflected or rotated versions of the target typically interferes with the search for a unique orientation in naïve observers (Zhaoping & Guyader, 2007). This interference, which might be the result of a top-down viewpoint-invariant shape recognition process, is manifested mainly by prolonged latencies to report the target's location after gazing at it first. We did not find an indication for a difference in bottom-up attentional processes between the groups, as eye-tracking revealed that the two groups gazed at the target comparably fast. There was not sufficient evidence to conclude that the ASD group differed from Controls in top-down processing. Nevertheless, we hope that our results will motivate future investigations.

3.2 Introduction

Autism is a developmental disorder whose mild to severe forms are collectively referred to as Autism Spectrum Disorders (ASDs). It is defined by behavioral criteria that fall into the categories of social and communication deficits and restricted and repetitive patterns of behavior and interests (APA, 2013). Despite impairments in these abilities, graphical and visuospatial skills are frequently found to be surprisingly well-developed.

3.2.1 Ignoring "global" context

In visual search the target is surrounded by distractor items which have to be ignored in order to find the target fast. People with autism outperform typical participants in conjunction search tasks and difficult feature searches (Plaisted et al., 1998; O'Riordan, 2000; O'Riordan, Plaisted et al., 2001; Iarocci et al., 2005; Jarrold et al., 2005; Baldassi et al., 2009; Joseph et al., 2009; Kaldy et al., 2011; O'Riordan, 2004; Kemner et al., 2008; Collignon et al., 2013; but see Keehn et al., 2013; Caron et al., 2006; Constable et al., 2010; Remington et al., 2009; Remington et al., 2012). It should be noted, however, that these findings are most robustly shown in children, while evidence for adults with autism is not yet abundant. There are two other visual tasks, the Embedded Figures Test (EFT), in which people with autism are frequently found to perform better compared with controls, and the Block Design Task (BDT), in which people with autism excel in comparison to their own performance on other cognitive tasks (Shah & Frith 1983; Shah & Frith 1993). The EFT requires spotting of simple geometric shapes embedded within more complex line drawings. In the BDT, arbitrary patterns have to be constructed from simple elements, by breaking down the larger pattern into its building blocks. In the EFT and BD task it is advantageous to ignore the larger pattern/the identity of the line drawing. Similarly, visual search is most efficient when distractors can be ignored. Common to all of these experimental paradigms is that processing of distracting, task-irrelevant visually presented information is detrimental to task performance (which does not imply, however, that findings of superiority in these tasks have the same explanation, cf. Mitchell and Ropar (2004)).

The Enhanced Perceptual Functioning (EPF) theory suggests that people with autism are able to ignore "higher-order" information if it is task-irrelevant, whereas processing of such information is mandatory in normal people, even if detrimental to the task (Mottron et al., 2006). At the same time, it is suggested by the EPF account that processing of higher-order information is possible in autism as long as it is relevant to the task (Mottron et al., 2003). By contrast, the Weak Central Coherence (WCC) account asserts that people with autism are unable to "integrate pieces of information into coherent wholes" (Happé & Booth, 2008, p.51) or at least do not tend to do so by default (Frith, 1989; Happé & Frith, 2006; Frith & Happé, 1994; Happé 1999). On first sight, it seems as if EPF and WCC yield different predictions for the processing of higher-order information. However, this is not necessarily the case, if they refer to distinct concepts by "global" or "higher-order" information. There are two reasons to believe that they do so; a historical ambiguity in the terms, and an indication for unrelatedness of the constructs that are used in EPF and WCC.

According to Mottron and Soulières (2013) the ambiguity that connotates the terms "local" and "global" has its origin in a detail-whole distinction made by Kanner in 1951, wherein the distinction had two meanings: (1) a within-perception meaning, where it denotes a hierarchical relation of levels within perceptual stimuli, and (2) a top-down meaning, where it refers to an atypical relationship between perception and higher-order cognitive processes, such as meaning of the perceived. For example, in (1) the global and local levels of a visual pattern can be distinguished by its relatively larger and relatively smaller spatial extent, such as the smaller and larger letters of a Navon hierarchical figure (Mottron & Soulières, 2013). In accordance with (2), the "overall meaning" of the line drawing in the EFT is what constitutes its higher-order level, by which people with autism are "captured" less compared with typical people (Shah & Frith, 1983). Although Shah and Frith (1993) used the terms local and global in connection with the then recent weak central coherence account, they must be understood in a top-down sense there. The EPF account, on the other hand, has used global and local processing in the first meaning. This has led to the understanding that the theories hold opposing positions toward the integrity of global processing in autism

(Mottron & Soulières, 2013). However, this is not the case, provided that the theories refer to different concepts by “global”.

Milne and Szczerbinski (2009) offer an exploratory factor analysis to better operationalize the concept of weak central coherence and to identify commonalities in 14 tasks that have been used to measure either central coherence or global-local perceptual style. They were motivated by the fact that the term “weak central coherence” has been used to imply a bias for local over global processing, although this is an extension of the original concept (Milne & Szczerbinski, 2009). Their results show that performance of 90 subjects of the normal (non-autistic) population is largely uncorrelated between the tasks. This suggests that these tasks do not measure a unitary construct. The only reliable factor of the analysis, which corresponded to central coherence, received substantial loading from the EFT and BD tests, confirming previous operationalization of the central coherence concept via performance on these two tests. Milne and Szczerbinski’s analysis also includes tasks that require global processing in Mottron et al.'s (2013) sense (i.e., the global level of a stimulus can be perceived independently of its local elements), such as Navon-style hierarchical figures. Those were found to be unrelated to the "weak central coherence" factor. Moreover, classic visual search (“red X within red T and green X letter search”), which has also been associated with central coherence in the literature (e.g., O’Riordan et al., 2001), did not load onto this factor. Milne and Szcerbinski (2009) suggest that central coherence should be regarded as unrelated to the capacity for integration, gestalt grouping or “global perception” (p. 18).

Of note is also the observation that findings of superior performance of people with ASD in the EFT can be interpreted according to the above distinction: they might be better able to (a) see parts of the whole visual stimulus or (b) ignore the meaning of the drawing (Mitchell & Ropar, 2004). Brian and Bryson (1996) set out to test these two hypotheses but failed to replicate the finding of superior performance of the autism group in the standard EFT. This study has since not been repeated to the best of our knowledge.

Together, the EPF account and WCC theory yield three predictions: People with autism can be assumed to be (a) able to process information on different levels of a visual stimulus (i.e., local and global level) equally well as typical people. These levels are determined by the relative spatial extent or quantity of information. People with autism are, in contrast to typical people, (b) able to ignore higher level perceptual information if it is irrelevant to the task. Moreover, they are (c) less likely to process the overall meaning of a stimulus, compared with typical people.

3.2.2 Experimental evidence for reduced top-down interference from task-irrelevant visual contextual information on a primary task

Contextual information, or information from top-down or prior knowledge, may sometimes interfere with the necessary processes to complete a primary task. Contextual information may refer to two levels here: Information of a visual context may be of a perceptual and semantic kind, such as in the EFT, where the global perceptual pattern depicts a meaningful scene at the same time. Alternatively, it may be of a perceptual kind only, as in some visual illusions, such as the Titchener illusion, in which it would be difficult to find a semantic interpretation of the visual pattern. Instead of classifying the diversity of visual tasks according to these categories (which might be impossible in some cases and would fall out of the scope of the current endeavor in any case), findings from experiments of both kinds will be briefly reviewed in the following.

Susceptibility to a variety of visual illusions of people with autism has been tested, but the findings are mixed (Titchener, Müller-Lyer, Ponzo, Poggendorff, Hering & Kaniza triangle used by Happé, 1996; Ponzo, Titchener, Müller-Lyer, Poggendorff and Hering illusions used by Bölte, Holtmann, Poustka, Scheurich, & Schmidt, 2007; Shepard illusion: Mitchell et al., 2010; but see Ropar & Mitchell, 1999; 2001). People with autism are better at copying impossible figures, although they are able to recognize their impossibility when asked, which may reflect their ability to ignore this information (Mottron, Belleville, & Ménard, 1999). In a recent study, Liu, Cherkassky, Minshew and Just (2011) hypothesized that people with ASD would show a reduced interference from impossible 3-D shapes on a lower-level line counting task. Their participants indicated whether there were more green or more red contours within 3-D shapes in one task, and judged the (im-) possibility of the same 3-D shapes in another task, while being scanned with functional magnetic resonance imaging. The groups were similar on all behavioral measures and in brain activation during the possibility judgments. However, whereas the control group showed increased activation in a medial frontal region in the line-counting task relative to the possibility task, this increase was absent in the autism group. This absence was interpreted as reduced effort in the counting task in ASD compared with the control group, attributed to reduced interference from the task-irrelevant 3D-object during that task. However, people with autism are not immune to influence by prior knowledge. For example, typical observers exaggerate the circularity of an oval shape when adjusting a test oval shape to match a standard. People with autism were less influenced by prior knowledge in this task when judging the oval shape in absence of contextual cues, but equally influenced as comparison participants when visual cues were present (Ropar & Mitchell, 2002).

3.2.3 Negative findings from top-down processing in visual search in autism

The following two studies tested top-down effects from the object level on visual search in autism with comparably simple stimuli to those used in the present experiment.

O’Riordan (2000) hypothesized that the superior ability of people with autism to discriminate targets from distractors—a prominent explanation of their faster RTs in visual search—might stem from a top-down modulation process that enhances signals of relevant over irrelevant stimuli. Priming may be considered such a form of top-down modulation on incoming stimuli. Priming effects could, thus, be enhanced in autism compared to controls. As previous research tended to find impaired kinds of top-down processing in autism (executive/control processes), this would not be a likely prediction, however. O'Riordan tested negative and positive object-based priming in conjunctive visual search tasks with colored letters. In the first experiment, the target (a red X) and distractor types (red T and green X) remained constant over trials in the "invariant" condition. In the "variant" condition, half of the distractors were always red but varied character across trials (any letter other than X), and half of them were always letters X but varied in color across trials. If inhibition of known distractors facilitates search in the invariant condition, this facilitation should be absent in the variant condition. If people with autism rely more on this top-down facilitation, their performance should deteriorate more in the variant condition, compared to controls. Contrary to this prediction, the autism group was affected equally by the manipulation of varying the distractors over trials, compared with the controls, but possibly slightly less than controls with respect to the most difficult trials (largest set size, target absent; see O'Riordan's figure 1). The second experiment tested the effect of positive object-based priming by comparing performance on three tasks. In the first two tasks, the target and distractors remained the same throughout the session (targets were a red X in task 1, and a green T in task 2, with red T and green X distractors in all three tasks). In the third task, the observers were told that the target could be either that of the first or that of the second task. Target detection could be facilitated in the first two, but not the third task, due to positive effects of target repetition from trial to trial. Priming effects were comparable between participant groups.

Remington et al. (2009) studied response interference from a distractor letter in visual search that was either the same as one of the possible target letters or a neutral letter. Target and distractors (1, 2, 4, or 6 total items) were arranged in an imaginary circle around the central fixation point. A slightly larger letter was shown left or right peripheral to the circle, which could be any letter but the possible targets (X or N) in the "neutral" condition, and was always the non-presented possible target in the "incompatible" condition. Observers had to press one of two keys to indicate which of the two targets was presented in the circle. Interference from the peripheral letter is present when RTs are longer in the incompatible condition compared to neutral. Adults with or without ASD did not differ in RTs or accuracy throughout the task.

3.2.4 Conclusion

People with autism seem to be less influenced negatively by task-irrelevant contextual information in a variety of tasks. This is evident from studies of the Embedded Figures Test, where they are better able to ignore the global, meaningful pattern while performing the primary detection task, and from more realistic judgments about stimuli in some visual illusions. However, two visual search studies that involved some form of top-down interference from letters did not show an advantage towards interference for people with ASD (O'Riordan, 2000; Remington et al., 2009).

3.2.5 Motivation

One motivation of this study is the finding of an absence or alleviation of interfering task-irrelevant top-down processing in people with autism in some visual tasks. An implication of the EPF account is that people with autism are able to ignore task-irrelevant perceptual information. Based on notions of this kind, we would expect people with autism to be facilitated in their performance compared to typical people in a task where task-irrelevant top-down processing is detrimental to performance.

The current study aimed to investigate two types of visual processes in autism: a fast, possibly low level, detection processes of a unique orientation feature, and a slower, probably higher level visual shape recognition process that is detrimental to the search task (Zhaoping & Guyader, 2007). Stimuli were arrays of many identical X or N shapes (see Figure 1 below). Within a stimulus, all shapes had an oblique bar that was uniformly tilted 45° to the left or to the right. Only the target shape contained a uniquely tilted bar, that could be 45° or 20° tilted to the opposite direction of the non-target shapes. Observers were asked to report via button press the location of the unique orientation (left or right half of stimulus). It has been suggested that lower level vision (i.e., V1) would suffice to solve the task (Li, 2002). Instead, higher level vision might be involved as well: The fact that the target and non-target shapes are identical, up to a rotation or reflection, "camouflages" the shape containing the uniquely tilted bar temporarily among distractor shapes (Zhaoping & Guyader, 2007). Often, one of the observers' first saccades goes to the target, but they subsequently search elsewhere (see Figure A1 in appendix). The best available guess for what happens subsequently is that a viewpoint-invariant shape recognition process interferes with the search for a unique feature (Zhaoping & Guyader, 2007). The time of first gaze arrival to the target in a trial was recorded via eye-tracking. Interference mainly manifests itself by a prolonged time to report the target location after gaze arrived to it. We would expect a reduction of the prolonged time to report the target in people with autism in this task compared with typical controls.

Furthermore, we tested whether the time of first gaze arrival differed between the two participant groups. Bertone, Mottron, Jelenic, and Faubert (2005) suggested that lateral inhibition in primary visual cortex in people with autism might be “enhanced”, leading to more narrow tuning curves of V1 neurons, and thus, enabling better discrimination of contrast-defined edge orientations in autism. If that were the case, perhaps an orientation feature would draw attention to it faster in autism than in typical people. We think there is not enough evidence to expect or preclude any differences in bottom-up control of attention between autism and typical people, and we therefore tested for this hypothesis as well.

3.3 Methods[2]

3.3.1 Participants

3.3.1.1 Number of participants

A group of 29 high-functioning adults with autism spectrum disorder and a control group of 30 observers completed the experiment. The data of two typical participants were excluded from analysis as they were non-naïve to the task. We also excluded data from 11 participants with ASD and from seven participants without ASD, some of which did not meet our eye-tracking quality criteria, and with some we could not perform eye-tracking at all (jointly referred to as "insufficiently tracked participants" or "participants with insufficient eye-tracking quality"). The ages of another three typical participants clearly exceeded the age range of the ASD group after we removed data of participants with insufficient eye-tracking quality (mean age of ASD group: 34.33 years, *SD* = 7.06; mean age of these three typical individuals: 57.7 years, *SD* = 1.15). We decided to exclude the data of these three oldest typical participants to avoid that possible differences between the groups were due to age-related effects. Characteristics of the remaining participants (18 ASD and 18 typical) are summarized in Table 2. Table B1 in the appendix shows characteristics of insufficiently tracked participants (11 ASD and 7 typical).

3.3.1.2 Sample statistics

In the following sections, "participants" will refer to the well-tracked participants only. Unless stated otherwise, a two-tailed, two-sample t-test has been applied with a significance level of $\alpha = 0.05$.

---

[2] Note that the sections stimuli, apparatus and procedure have been reported previously in a bachelor's thesis (Tribull, 2011). The sections data analysis and results include data from 59 participants, data of 38 of which were obtained prior to commencing this MPhil thesis. Results of a preliminary data analysis with data of sufficiently high eye-tracking quality from 23 out of the 38 participants were reported in the bachelor's thesis.

Participants with ASD and those without did not differ in terms of chronological age, $t(34) = -0.53$, $p = .60$, verbal IQ, $t(33) = 1.32$, $p = .19$, performance IQ, $t(34) = 0.17$, $p = .87$, or Block Design score, $t(32) = -0.09$, $p = .93$, of the Wechsler Adult Intelligence Scale. Participants with ASD all had provided proof of an expert clinical diagnosis and, in addition, were tested on the Autism Diagnostic Observation Schedule (ADOS, Lord et al., 2000). In addition, an autism spectrum quotient (AQ, Baron-Cohen et al., 2001) had been obtained for 33/36 participants. Participants with ASD and Control participants differed in AQ, as would be expected, $t(31) = 6.93$, $p < .001$.

3.3.1.3 Other participant characteristics

All participants had normal or corrected-to-normal vision and were naïve to the experiment. Participants gave written informed consent at the beginning of the experimental session and received £ 10.00 at the end. This study was approved by the UCL Research Ethics Committee.

Table 2

Characteristics of participants (N = 36)

Group	ASD (*N* = 18)			Control (*N* = 18)		
Female : male	8 : 10			6 : 12		
	M	SD	Range	M	SD	Range
Age (years)	34.33	7.06	25 - 50	35.67	8.08	24 - 49
vIQ[a]	116.39	13.1	90 - 140	110.88	11.38	90 - 128
pIQ	110.56	17.62	77 - 138	109.72	11.7	79 - 131
AQ[b] *	34.65	10.39	11 - 47	14.81	4.96	7 - 28
ADOS score	9.17	3.91	3 - 17	-	-	-
BD score[c]	12.63	3.38	6 - 18	12.72	2.63	9 - 18

ASD Autism Spectrum Disorder, *N* number of subjects, *M* mean, *SD* standard deviation, *vIQ* verbal intelligence quotient, *pIQ* performance intelligence quotient, *AQ* autism quotient, *ADOS* autism diagnostic observation schedule, *BD* Block Design

[a] *N*(Control) = 17

[b] *N*(ASD) = 17; *N*(Control) = 16

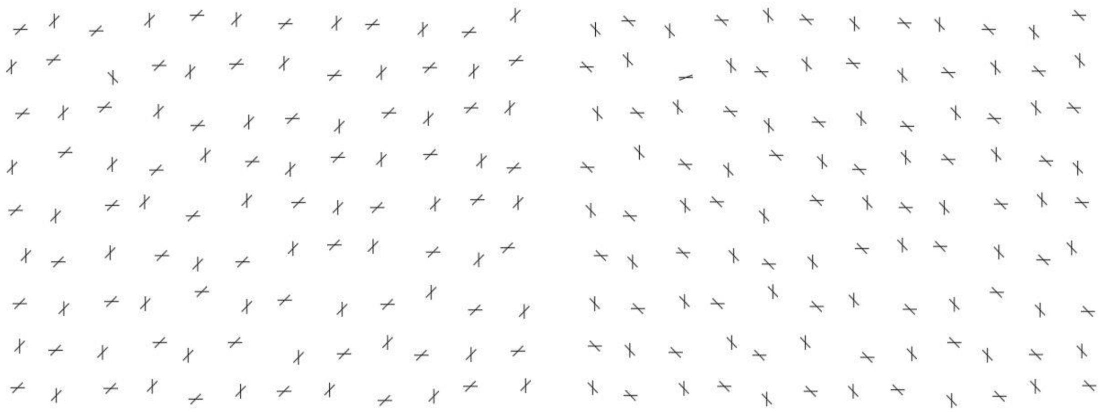
[c] *N*(ASD) = 16

**p* < .001 by unpaired, two-tailed t-test

3.3.2 Stimuli

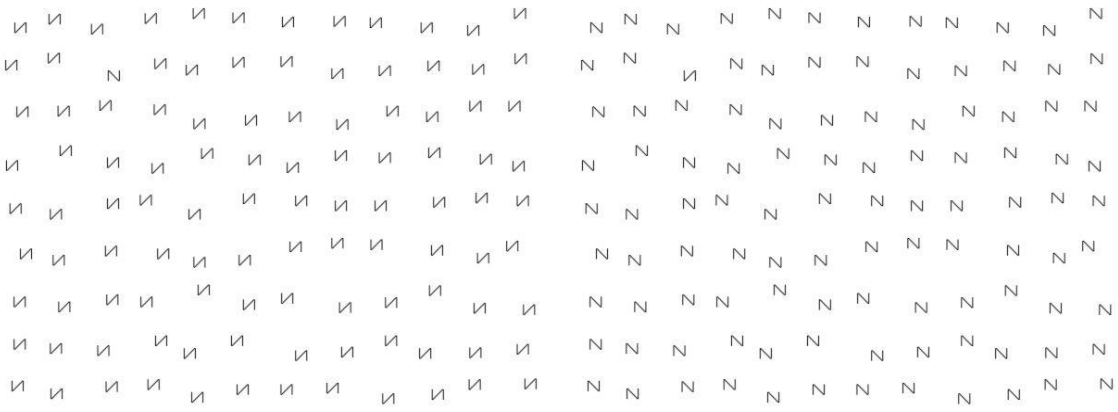
Four experimental conditions were defined by target shape within the search display (Figure 1). These were very similar to the stimuli used in (Zhaoping & Guyader, 2007), and (Zhaoping & Frith, 2011). In the X-search condition, a bar that was 45° left- or right tilted and a superimposed horizontal or vertical bar formed the target shape, while distractors were mirror-reflected or 90°-rotated versions of the target shape. The SX-search condition, sharing its distractor types with X-search, contained a “skinny” version of the X-shape as a target, with an oblique bar that was 20° left- or right tilted from horizontal or vertical. The 20° tilt gives the target of SX-search a unique shape within the stimulus, so that the object shape formed by the two intersecting bars is not perceived as identical to the distractors. SX-search does not give rise to interference between viewpoint invariant object recognition and the process of locating the target bar (Zhaoping & Guyader, 2007). The other two conditions N-search and RN-search were an N-shaped target within its mirror-reversal as distractors, and a mirror-reversed N-shaped target within normal N-shape distractors, respectively (Figure 1 C & D). The fact that it takes observers longer to search for the N-shaped target within its mirror-reversals than to search for the mirror-reversal among normal N-shapes (the reversed letter effect) was established by Frith (1974). Participants performed a single experimental session comprising 240 trials, in which trials of all four search conditions were randomly interleaved (60 trials per condition).

The stimulus displays were 600 pixels x 800 pixels (height x width) and spanned 35° x 46° of visual angle, viewed from a distance of 40 cm. Items consisted of two or three black 21-pixel-long and 2-pixels-wide bars presented on a uniformly grey background. Items were displaced—“jittered”—up to 15 pixels horizontally and/or vertically from their position in a regular grid of 9 x 12 rows by columns. In each trial, the target location was chosen at random from 16 possible locations on a circle with a radius of 272 pixels around the display center (± jitter), avoiding a region of 124 pixels to both sides horizontally from the display center. The fixation stimulus was a centrally displayed black disk of size 0.3° of visual angle on a grey (level intensity127) background.



A. X-search: X-shape in rotated versions of itself

B. SX-search: “skinny” X in distractors of X-search



C. N-search: N-shape in mirror-reversed versions of itself

D. RN-search: mirror-reversed N-shape in normal N-shape

Figure 1. Illustrations of the experimental search stimuli. The stimuli contained 9 rows by 12 columns of items. Note that the target is in the upper left region in all examples. A and B contain the same set of distractors (i.e. non-targets) but differ in their target types. A gives rise to shape-to-feature interference, whereas B does not, since target and distractors are reflected or rotated versions of each other in A, whereas in B target and distractors differ in shape. There were four possible appearances of targets and distractors in X-search and SX-search (as the target shape could be a combination of a right or left tilted target bar and a horizontal or vertical dissecting bar). C evokes stronger shape-to-feature interference than D.

3.3.3 Apparatus

The experiment ran on a Dell computer using a custom-written Matlab program under Windows XP. Displays were presented on a CRT screen. Gaze was tracked with a 50 Hz infrared video tracker from Cambridge Research Systems (www.crsltd.com). Calibration was performed at the beginning of the session before the test trials with a typical precision of 0.5°. Observers placed their chin on a chinrest during trials in a distance of 40 cm to the screen. They gave button responses to keys on a USB-connected numeric keypad, either using two fingers of the same hand or each thumb for the designated keys. The experiment was programmed and all button presses and the eye-tracking data recorded using Matlab and the Matlab Psychophysics Toolbox.

3.3.4 Procedure

Observers started each trial by a button press, which was followed by presentation of the fixation stimulus. They were instructed to fixate the central disk until it disappeared, and to move their eyes freely to search for the target thereafter. If no eye-tracking was performed, the fixation stimulus disappeared after 600 ms, otherwise it remained until gaze stayed in a radius of 32 mm around the disk for at least 40 ms. A blank (uniformly grey) screen replaced the fixation stimulus for 200 ms before the onset of the search stimulus. Observers were instructed to report the target location (left or right side of the display) via the appropriate key press as soon and as accurately as possible. Once a button response was obtained the search stimulus was replaced after 300 ms by a display instructing to “press any button to continue” (see Figure 2). Gaze was tracked from the onset of the fixation stimulus until the presentation of this instruction.

The target was defined as “the item containing a uniquely tilted bar”. Observers were told that any vertical or horizontal bars within the stimulus image could be ignored, and that the uniquely tilted bar could be oriented to the left or to the right at random in each trial. We avoided reference to the search items as “letters” or “shapes” throughout the session, but instead referred to them as "items", which is common in this type of study, to avoid possible priming for attention to shapes. In order to prevent systematic search, we told observers to "avoid scanning the screen line by line as in reading”.

Participants performed 16 practice trials during the task instructions, data of which were not analyzed. Four trials of each condition were randomly interleaved during practice, covering all possible target shapes and orientations. Observers received auditory feedback for correct/incorrect button presses or the pressing of non-designated keys during practice but not during the 240 experimental trials. They were given a short break after every 80 trials (twice in a session).

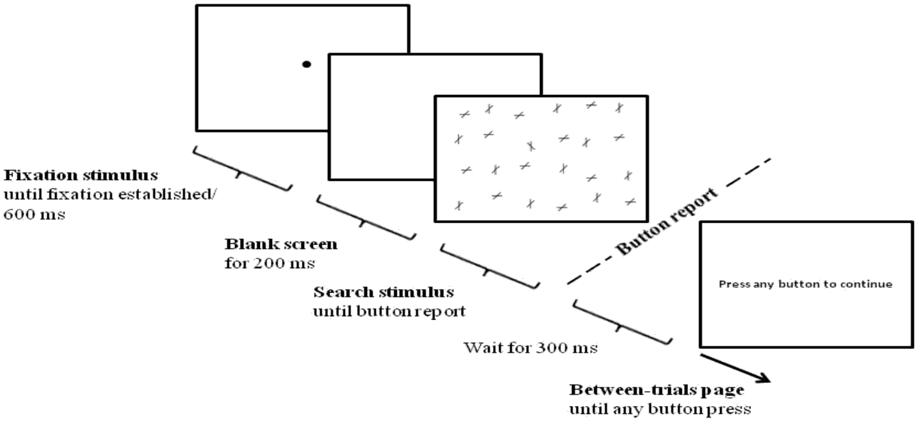


Figure 2. Sequence of events from one trial to the next. Note that experimental stimuli contained more items than shown in this schematic.

3.3.5 Data-Analysis

We decided to conduct several ANOVAs in order to test for differences between the mean reaction times of the ASD group and the Control group, with respect to the kind of reaction time measure and experimental condition. ANOVA has an advantage over multiple applications of the two-sample t-test, as the chance for type-I errors (false rejections of the null hypothesis) is not increased for the same number of comparisons. This test makes specific assumptions about the underlying distributions of the data, among which is the assumption of normally distributed data. Although our data deviated somewhat from a normal distribution, we decided to conduct the analysis nevertheless, because ANOVA is relatively robust against violations of this assumption (Schmider, Ziegler, Danay, Beyer & Bühner, 2010).

3.3.5.1 Symbols

We denote by RT_{report} the time of the observer’s button press after stimulus onset. RT_{eye} is the time when gaze arrived at the target for the first time within a trial. RT_{lapse} is the time lapse between the time of gaze arrival RT_{eye} and the time of button press RT_{report}, so that:

$$RT_{lapse} = RT_{report} - RT_{eye}.$$

3.3.5.2 Eye-tracking

We defined “gaze arrival” as the first time within a trial when gaze was detected within a radius of r = 22.6 mm around the central pixel of the target. This radius was defined as a trade-off between two competing criteria: on the one hand it must not be too small, as eye-tracking error (up to 0.25° of visual angle for our equipment) and errors in the landing of saccades have to be taken into account. On the other hand, it should not be so large as to encompass several of the search items—if it did, we might falsely count the looking at a non-target item contained within the radius as a gaze arrival to the target. A range of values would fulfill this trade-off. Initially, we chose for the gaze arrival criterion the length of 0.7 times the distance between items on a regular grid (= 22.6 mm). In order to make sure that our results did not depend on this particular choice, we also analyzed data with respect to several other choices of gaze criteria (see section 3.5.2).

Obviously, data from trials in which gaze never arrived at the target (“non-arrival trials”) could not be divided into RT_{eye} and RT_{lapse}, and hence could not be interpreted with respect to the underlying attentional processes. Non-arrival trials do not necessarily mean that the observer was guessing the answer. Some observers note the presence of a salient target in the periphery of the visual field, so that they do not have to look directly at the target. Thus, they may saccade toward it, but not close enough to be within our gaze arrival criteria.

3.3.5.3 Trial inclusion

Results presented below contain data from "valid" trials only. A trial was defined as “good” if less than 10% of frames between the time of stimulus onset and RT_{report} were untracked, and was defined as “bad” otherwise. A trial was defined as “valid” if it was a good trial, and, in addition, gaze arrived at the target and observers gave a correct button response at least 100 ms after search stimulus onset.

3.3.5.4 Response logging

A correct response required the observer to press the designated left button (key “4”) when the target was on the left half of the display, and the right button (key “+”) when the target was within the right half of the display. The target was present in all trials. The stimuli were created in a way such that the target could never appear in an ambiguous position in the middle between the left and the right half of the display.

3.3.5.5 Participant inclusion criteria

We only analyzed data of participants with sufficiently high overall eye-tracking quality. Participants were said to have sufficiently high overall tracking quality if more than 80% of trials out of the total 240 trials were good trials.

3.4 Results

3.4.1 Trial-exclusion statistics

The following trial rates were calculated for each subject within each condition as n/60, where n is the number of trials with a certain attribute (e.g. non-arrival trials) and 60 is the total number of trials in a given condition. Note that a given trial could contribute to more than one of the rates. Figure C1 in the appendix shows all individual and mean rates.

3.4.1.1 Trials with RT_{report} < 100 ms

In the ASD group there were 3 trials out of 240 x 18 trials and in the Control group 4 trials out of 240 x 18 trials in which RT_{report} < 100 ms. Thus, these trials only formed a negligible proportion of excluded trials.

3.4.1.2 Rates of non-arrival trials

In the ASD group, the highest rates of non-arrival trials were 25% (one subject in X-search, one subject in SX-search) and 17% (one subject in N-search). There were another four occurrences of non-arrival rates between 6% and 10% (1 x N-search, 3 x RN-search). All remaining individual non-arrival rates as well as the means were below 5% in this group.

In the Control group, the highest non-arrival rates were 27% and 15% (both in RN-search), and 7% (one subject in SX-search). All remaining individual as well as the mean non-arrival rates were below 5% in this group.

3.4.1.3 Rates of trials with insufficient eye-tracking quality (bad trials)

In the ASD group, the highest rate of bad trials was 27% (two participants, RN-search and SX-search, respectively). There were four occurrences in X-search, another two occurrences in SX-search, and one in N-search, of bad trial rates between 16% and 25% in this group. All other individual rates of bad trials were below 15% in this group, with all mean rates below 10%.

In the Control group, the highest rate of bad trials was 27% (one participant, X-search). Six participants had bad trial rates between 10% and 20% (2 x N-search, 3 x X-search, 1 x SX-search). All remaining individual as well as mean rates of bad trials were below 10% in this group.

3.4.1.4 Error rates

Error rates were calculated as the number of trials with an incorrect button press divided by the total number of trials within a condition. In the ASD group the highest error rates were 34%, and 24% (1 x RN-search, 1 x X-search). Another seven rates were between 10% and 16% (5 x RN-search, 1 x X-search, 1 x SX-search). All other individual error rates were below 10% in this group, with all mean error rates below 5%.

In the Control group, the highest error rates were 34% and 17% (3 and 6 participants, respectively, all in N-search), and 12% (1 x X-search). All other individual as well as mean error rates were below 10% in this group.

3.4.1.5 Proportion of valid trials

The mean proportion of valid trials among all trials in the ASD, *M* = 0.89, *SD* = 0.07, and in the Control group, *M* = 0.92, *SD* = 0.04, did not differ significantly, *t*(34) = -1.3, *p* = .20. This proportion was also compared between the groups within each condition separately, yielding no significant differences (see appendix C).

The lowest numbers of valid trials among participants and conditions were 40 and 42 in the ASD and in the Control group, respectively (in X-search in both groups). Thus, every participant had at least 40/60 (67%) trials in each condition to be included in the analysis.

3.4.2 On average over trials

3.4.2.1 Response times

An ANOVA was conducted on RTs including all valid trials with group (ASD, Control) as between-subject factor, and condition (N-search, RN-search, X-search, SX-search) and RT-type (RT_{eye}, RT_{lapse}) as within-subject factors.

The ANOVA showed significant main effects of RT-type, *F*(1,34) = 8.69, *p* = .006, and of condition, *F*(3,34) = 36.23, *p* < .001, but not of group, *F*(1,34) = 0.13, *p* = .72. The interaction of RT-type and condition was significant, *F*(3,34) = 25.31, *p* < .001. The other three interactions were not significant (group x RT-type: *F*(1,34) = 0.34, *p* = .57; group x condition: *F*(3,34) = 0.18, *p* = .91; group x RT-type x condition: *F*(3,34) = 0.77, *p* = .51). Hence, there were no significant differences between the groups' RTs on average over trials.

3.4.2.2 Accuracy

Although the proportion of excluded trials did not differ between the groups, it could still be possible that the groups differed in their accuracies in one or more conditions. This aspect is important since the task can give rise to a speed-accuracy tradeoff.

An ANOVA was conducted on the error rates with group (ASD, Control) as between-subjects factor and condition (N-search, RN-search, X-search, SX-search) as within-subject factor.

The ANOVA showed a significant main effect of search condition, *F*(3,34) = 10.37, *p* < .001. The other main effect of group and the interaction were not significant (group: *F*(1,34) = 0.08, *p* = .77; group x condition: *F*(3,34) = 0.76, *p* = .52). Thus, the groups did not differ significantly with respect to their error rates on average over trials.

3.4.3 Initial trials

A previous study found that normal observers had high RTs in the beginning of X-search but quickly learned to reduce them, such that application of right-parietal repetitive transcranial magnetic stimulation became ineffective in this condition already after the first 10 trials (Oliveri et al., 2010). In the study of Oliveri et al., observers were given 7 trials for training (comprising both X-search and SX-search) before the start of the experimental trials. Here, observers were given more than twice as much training (16 trials; although with four different conditions interleaved instead of two). We hypothesized, therefore, that observers might complete learning even sooner than within the first 10 trials, and decided to analyze RTs within the first 5 trials of each condition, in addition to the analysis of RTs on average over all trials.

As displays of the four different search conditions were randomly interleaved throughout the experiment, the first trial of a given condition might have appeared as the second, third, fourth, etc. trial within the sequence of all trials. Hence, “trials 1-5” refers to the first 5 trials of a given condition, as opposed to the first 5 trials in the overall sequence of experimental trials.

For each of the conditions, we selected the first 5 trials and then only selected those that were valid among these 5 trials to be included in the analysis. We repeated the same ANOVAs for the initial 5 trials of each condition as were conducted for all trials.

3.4.3.1 Proportion of valid initial trials

Each participant had at least 2 valid trials and at most 2 incorrect trials out of the first 5 trials in each condition. The mean proportion of valid trials out of the first 5 trials per condition did not differ significantly between the groups with respect to any condition (N-search: $t(34) = -0.43$, $p = .67$; RN-search: $t(34) = -0.20$, $p = .84$; X-search: $t(34) = 0.77$, $p = .43$; SX-search: $t(34) = 0.35$, $p = .73$).

3.4.3.2 Indices of valid initial trials

As trials of each condition were randomly interleaved, each participant experienced a different sequence of search types. Therefore, it would be hypothetically possible that participants of one group experienced on average the first 5 trials of a given condition later than the other group. Furthermore, the exclusion of invalid trials might render the indices of trials selected for analysis unequal between the two groups. Both of these factors would lead to an unfair comparison of RTs between the groups within these initial 5 trials. The distributions of indices of the first 5 valid trials within a given condition were compared between the two groups with a two-sample Kolmogorov-Smirnov test, which tests whether two datasets stem from the same underlying distribution. The test yielded no significant differences with respect to search condition (N-search: $k = 0.15$, $p = .29$; RN-search: $k = 0.14$, $p = .38$; X-search: $k = 0.10$, $p = .76$;

SX-search: $k = 0.13$, $p = .46$). As the distributions of trial indices did not differ significantly between the groups in any condition, it can be assumed that the selected trials represent comparable time periods in both groups.

3.4.3.3 Response times

The ANOVA for response times in initial trials had the same design as the ANOVA over all trials, i.e., group x RT-type x condition, where RT-type and condition were within-subjects factors of repeated measures, and group was a between-subjects factor. It showed no significant main effect for group, $F(1,34) = 3.17$, $p = .08$, although this p-value is close to significance. The other two main effects were significant (RT-type: $F(1,34) = 9.7$, $p = .004$; condition: $F(3,34) = 23.51$, $p < .001$), and so were all interactions (group x RT-type: $F(1,34) = 8.39$, $p = .007$; group x condition: $F(3,34) = 4.06$, $p = .009$; RT-type x condition: $F(3,34) = 12.69$, $p < .001$; group x RT-type x condition: $F(3,34) = 10.17$, $p < .001$).

First, two separate follow-up ANOVAs with condition as within-subjects factor and group as between subjects-factor were conducted for RT_{eye} and RT_{lapse}. RT_{eye} was not significantly different between the groups with respect to any search condition (the main effect of group was not significant, $F(1,34) = 0.57$, $p = .46$, and neither was the interaction of group x condition, $F(3,34) = 1.41$, $p = .24$). The follow-up ANOVA with RT_{lapse} as dependent variable showed a significant main effect of group, $F(1,34) = 5.66$, $p = .02$, and a significant interaction of group x condition, $F(3,34) = 7.63$, $p = .0001$. Importantly, the interaction effect was due to a significant difference between the groups in their RT_{lapse} in the first 5 trials of X-search only, $t(34) = 2.86$, $p = .007$, as can be seen in Figure 3 panel A.The groups did not differ significantly in RT_{lapse} in the other three search types (N-search: $t(34) = 1.38$, $p = .18$; RN-search: $t(34) = -0.08$, $p = .94$; SX-search: $t(34) = 0.29$, $p = .77$).

3.4.3.4 Accuracy

The ANOVA on error rates within the first 5 trials of each condition did not show any significant effects (group: $F(1,34) = 0.35$, $p = .56$; condition: $F(3,34) = 2.25$, $p = .09$; group x condition: $F(3,34) = 0.25$, $p = .86$). However, when inspecting Figure 3 panel C, it looks like there is a trend for a higher error rate of ASD subjects in X-search. Comparing the two groups' accuracy rates in X-search trials 1-5 with a t-test turned out insignificant, $t(34) = 1.14$, $p = .26$.

We also conducted a Bayes factor analysis of this insignificant result to test for the possibility of a speed-accuracy tradeoff. In particular, if there was a speed-accuracy tradeoff, the ASD group should have a higher error rate than Controls in X-search trials 1-5. Bayes factors (*B*) indicate the relative strength of evidence for two theories (Dienes, 2014).

Parameters of the analysis are presented in Table 3. We used a free online calculator (http://www.lifesci.sussex.ac.uk/home/Zoltan_Dienes/inference/bayes_factor.swf). Error rates were translated into accuracy rates for better handling in the Bayes factor analysis (i.e., accuracy rate = 100 – error rate*100). A uniform distribution was chosen instead of half normal or normal because there was no reason to expect a particular accuracy value from the ASD group. If, for example, a difference between the groups in accuracy rates of 20% could be expected, a normal distribution centered at 20% should be used. Because of our restriction that subjects included in the analysis could have no more than 2 trials with incorrect button reports within the first 5 trials, the lowest possible accuracy rate was 60% (for 3/5 trials with correct button reports). Testing for the hypothesis that the ASD group has a lower accuracy rate than Controls gives an upper bound of the Control mean accuracy rate of 96.67 %. Thus, the centered bounds were 0 and 36.67, for lower and upper bound, respectively.

According to Dienes (2014), a Bayes factor close to zero implies strong evidence for the null hypothesis, close to 1 supports neither hypothesis, whereas much greater than 1 means strong evidence for the alternative hypothesis. Recommended cut-offs are at 1/3 and 3, respectively, for the three possibilities. The *B* obtained for trials 1-5 was 0.33, which is exactly the recommended cut-off between strong evidence for the null hypothesis and insensitivity. We also tested the same hypothesis using trials 1-10 instead (see bottom row of Table 3). The resulting *B* of 1.17 implies that neither hypothesis is supported. Thus, there is not sufficient evidence to assume a speed-accuracy tradeoff in the ASD group.

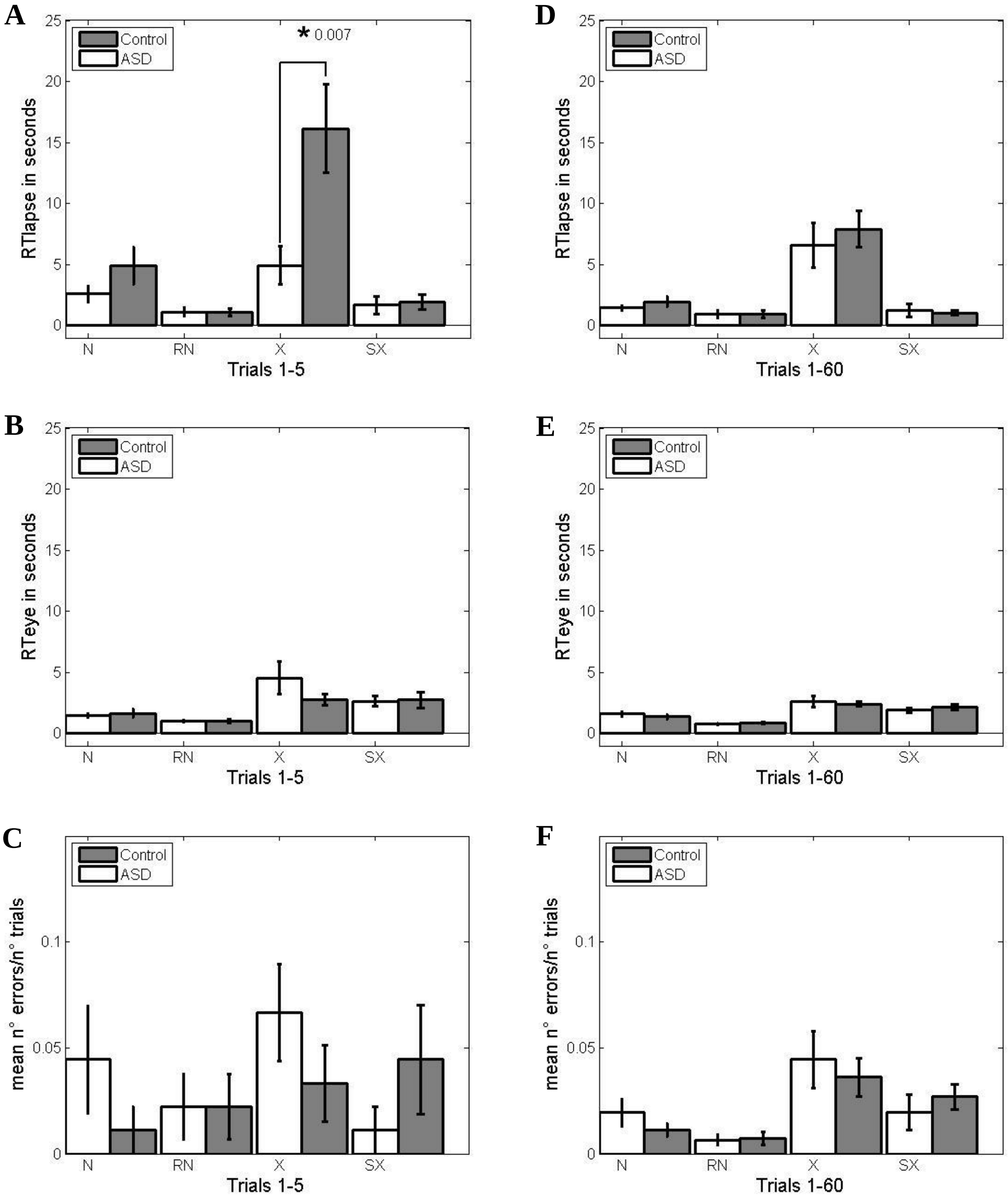


Figure 3. Mean response times and error rates. A-C Response times and error rates within the first 5 trials of each search type (N, RN, X, SX). D-F Response times and error rates averaged over all 60 trials of each search type. In each of the panels A,B,D, and E, the bars show averaged response times (A & D: RT_{lapse}, B & E: RT_{eye}) over valid trials within a search condition. Error bars show standard error of the mean. The asterisk and p-value in panel A show a significant difference between the groups in trials 1-5 of X-search (two-tailed unpaired t-test, $p = .007$). C & F: Error rates were calculated as the proportion of trials with an incorrect button report out of the first 5 trials (panel C) or out of all trials (panel F) within a given condition.

Table 3

Parameters and results of Bayes factor analysis

X-search trials	ASD mean accuracy rate (% correct trials)	Control mean accuracy rate (% correct trials)	Difference (effect size)	*p*	*t*	SE	Uniform distribution bounds	*B*
1-5	93.33	96.67	3.33	0.26	-1.14	-2.91	ASD accuracy is expected btw 60% and 96.67%. Centered bounds: [0, 36.67]	0.33
1-10	91.67	97.22	5.55	0.06	-1.90	-2.92	ASD accuracy is expected btw 60% and 97.22%. Centered bounds: [0, 37.22]	1.17

ASD autism spectrum disorder, *p* p-value, *t* t-value, *SE* standard error, *B* Bayes factor.

Note. Hypothesis: Accuracy rate of ASD is lower than that of Control. Assumption: Accuracy cannot be below 60%.

3.5 Robustness of results

It is also possible that the single significant result was obtained by pure chance. Quite a few conditions were tested, and it is quite likely to find a significant difference by pure chance the more tests are performed. The expected number of significant differences by pure chance (false discoveries) if all 24 comparisons between the two groups as presented in Figure 3 were performed (assuming that these were not guided by ANOVA outcomes) is 24*0.05 = 1.2; thus, the one significant t-test outcome might be a false rejection of the null hypothesis. The finding would be more convincing if it did not depend on a contingent parameter setting of our data analysis. In particular, we tested whether the result changed qualitatively with manipulation of two parameters that are in principle up to the researcher to decide, although there were reasons for our particular choice. These are the particular choice of interval of initial trials and the gaze arrival criterion. We compared ASD and Control groups after varying these parameters and arrived at the same conclusion as with initial parameter settings. We also tested whether the result would be the same when using a non-parametric hypothesis test (Mann-Whitney U-test/Wilcoxon rank sum test), which also turned out significant.

There are statistical methods to control for the family wise error rate (FWER), i.e. the probability of making one or more false discoveries. If we apply the most conservative common correction for multiple hypothesis testing, the Bonferroni correction, the adjusted significance level for the total of comparisons becomes 0.05/24 = 0.002. According to this correction, the p-value of 0.007 from the comparison between groups in RT_{lapse} of X-search trials 1-5 must be regarded insignificant. However, these methods typically assume individual tests to be independent, which is not given in our case when comparing several trial intervals, some of which overlap, so that the Bonferroni correction might in fact overestimate the FWER here.

3.5.1 Varying limits of the interval of initial trials

One could imagine that the groups cede to differ when tested on trials 1-6, 1-10, or 2-6 etc. When looking at Figure 4, it is obvious that several choices of trial interval should yield qualitatively the same conclusion, although the high fluctuation in Control RT_{lapse} in initial trials will probably render some comparisons insignificant. Varying the upper trial limit did not render the difference insignificant for several choices of trial periods whereas varying the lower trial limit did influence the significance somewhat (see Figure 5).

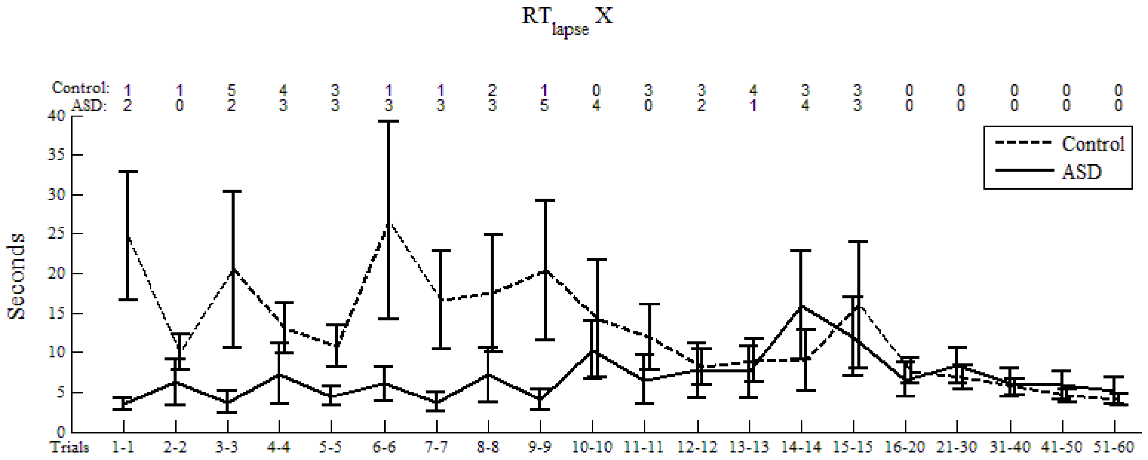
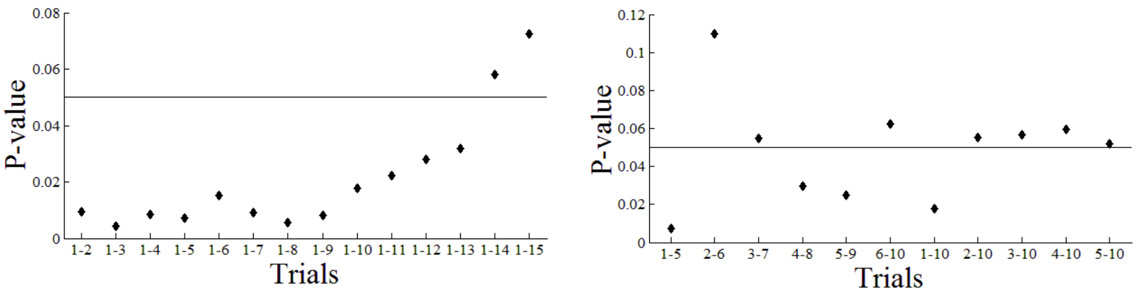


Figure 4. Mean RT_{lapse} in X-search for several trial periods. Errorbars show standard error of the mean. The lines connecting error bars show the mean RT_{lapse} within each trial period indicated on the x-axis. The two top rows of numbers show how many individuals did not contribute a mean value in a given trial period (due to too many invalid trials). Upper row: control group; lower row: ASD group. Missing values were ignored when averaging over subjects within a group.



A: Varying upper trial limit

B: Varying lower trial limit

Figure 5. P-values of t-test between groups for varying intervals of trials. The dots are the p-value of a two-tailed unpaired t-test of RT_{lapse} in X-search between the groups, for each of the periods of trials indicated on the x-axis. The line shows the 0.05 significance mark. A: The significance of the group difference persists with several upper trial limits within the initial trials. B: Varying the lower trial limit does affect the significance somewhat.

3.5.2 Varying gaze criterion

Our default criterion for gaze arrival at the target was a radius of 46 pixels (22.6 mm) around the target. Reducing or increasing this distance did not affect the significance of the result (see Figure 6). Explanations for the choices of gaze criteria are given in Table D1 in the appendix.

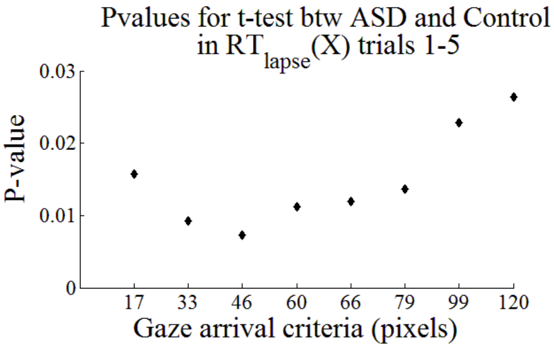


Figure 6. Change of p-values for increased/decreased gaze criteria. RT_{lapse} in trials 1-5 of X-search was compared between the groups for several choices of gaze criteria with a two-tailed unpaired t-test, p-values of which are shown here. No reasonable change of gaze criteria renders the group difference insignificant.

3.5.3 Using a different hypothesis test

Student's t-test was used for comparisons between the groups with respect to RT_{lapse} in each of the search conditions separately. This test compares two means. However, the mean is vulnerable to extreme values. By looking at the 36 individual RT_{lapse} within trials 1-5 (Figure 7A), we were concerned that the mean could have been drawn upward by a few high individual values. We, therefore, compared the two groups' medians of RT_{lapse}, which is robust against extreme values, using a Wilcoxon rank sum test. This is a nonparametric test that does not, unlike the two-sample t-test, assume the variances of the two samples to be equal. This test also turned out significant for trials 1-5 (*p* = .003).

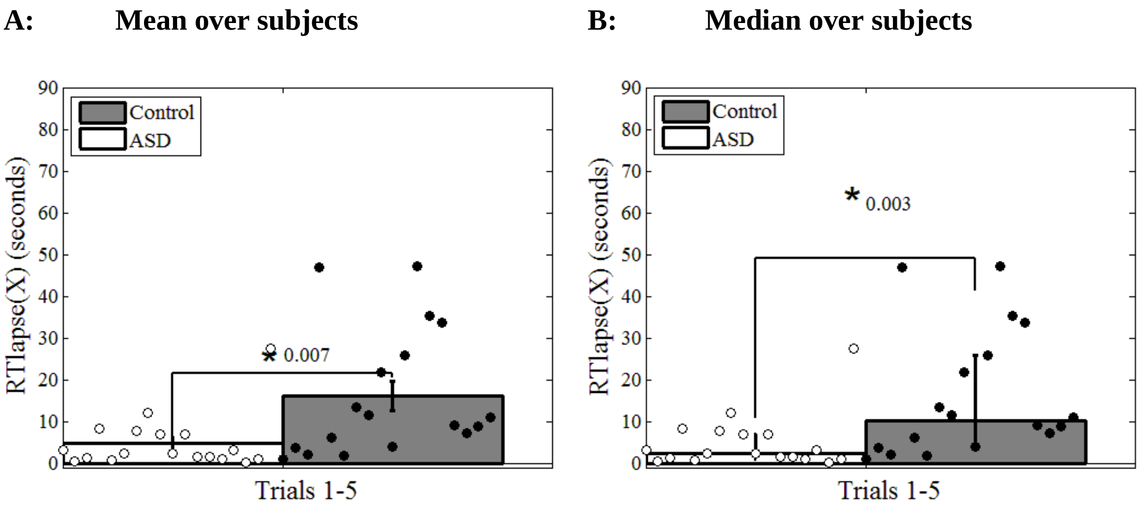


Figure 7. Individual RT_{lapse} and their mean or median within trials 1-5 of X-search. Dots are data of individual subjects. The bars show mean RT_{lapse} (panel A) and median RT_{lapse} (panel B) within each group. A: The asterisk indicates a significant difference between groups by a two-tailed unpaired t-test (*p* = .007). Error bars show standard error of the mean. B: The asterisk indicates a significant difference between the groups by a Wilcoxon rank sum test (*p* = .003). Error bars show the 25th and 75th quartiles of the data.

3.5.4 Comparison between the well-tracked and the insufficiently tracked participants

In order to reduce the suspicion that exclusion of participants with insufficient eye-tracking quality resulted in any bias in the sample of participants that were selected for data analysis, we compared well-tracked and insufficiently tracked participants within each group (ASD, Control) on the basis of their pure button report times. Note that we used all trials with correct button presses for this analysis, including trials with poor eye-data (bad trials, see section 3.3.5.3). Table 4 reveals no difference in RT_{report} (mean and standard deviation) between the well-tracked and the insufficiently tracked Control participants by t-tests between these two subgroups for trials 1-5 in each condition. However, there was a significant difference (p = .02) between well-tracked and insufficiently tracked ASD participants: Table 5 shows RT_{report} (mean and standard deviation) of the well-tracked and the insufficiently tracked ASD participants as well as the results of t-tests between these two subgroups for trials 1-5 in each condition. Marked in gray is a significant difference, i.e. the well-tracked ASD subgroup has significantly lower RT_{report} in X-search than the insufficiently tracked ASD subgroup. Moreover, the excluded ASD participants have an RT_{report} that is even higher than that of the well-tracked control group on average over subjects. There were no significant differences between the well-tracked and insufficiently tracked subgroups within the ASD group in chronological age, $t(27) = -1.7$, $p = .10$, verbal IQ, $t(27) = -0.05$, $p = .96$, performance IQ, $t(27) = 0.56$, $p = .58$, AQ, $t(27) = -1.42$, $p = .17$, ADOS, $t(27) = -0.54$, $p = .59$, or BD score, $t(27) = -0.58$, $p = .57$. Mean error rates for X-search compared between well-tracked and insufficiently tracked ASD participants did not differ significantly, $t(27) = .74$, $p = .46$. However, note that the insufficiently tracked sample contains 11 subjects, out of which only one (9%) made an error within the first five X-search trials, whereas 6/18 (33%) well-tracked ASD participants erred in one out of the first five trials (see Figure E1 in the appendix). But, as we discussed above, there is not enough evidence for the hypothesis that the well-tracked ASD sample achieved lower RTs by being less cautious with reporting the target.

There is reason to worry that the sample used for the analysis with eye-data is biased. Note that insufficient eye-data prevents us from knowing whether differences between groups in RT_{report} may result from differences in RT_{lapse} or RT_{eye}. Having found no differences between well-tracked ASD and Controls in RT_{eye}, however, we may assume that differences between the groups in RT_{report} are more likely to result from differences in RT_{lapse}. Thus, it is possible that the true ASD population mean RT_{lapse} is closer to that of typical people than is estimated with our well-tracked sample. This can also be seen in Figure 8, which shows RT_{report} for all naïve subjects regardless of tracking quality. In contrast to eye-data, RT_{report} was obtained for all subjects who participated in the experiment. Although RT_{report} does not solely depend on RT_{lapse}, a difference between the groups in RT_{lapse} of several seconds should reflect in RT_{report} as well. When examining Figure 8, however, it might be possible that any difference between groups resulted from random fluctuations. This is consistent with the hypothesis that the well-tracked subgroup of ASD subjects is biased towards lower RT_{lapse} compared to the true ASD population mean RT_{lapse}. Alternatively, the excluded ASD subgroup might have had higher RT_{report} (and supposedly RT_{lapse}) compared to well-tracked participants for a reason that was related to the cause for insufficient eye-tracking quality; however, the answer remains unknown.

Table 4

Comparison between well-tracked and insufficiently tracked Control participants' RT_{report} in trials 1-5

Group		Well-tracked Control (*N* = 18)		Insuff. tracked Control (*N* = 7)			
		M	SD	M	SD	*t*	*p*
RT_{report}	N	6.719	7.203	2.882	1.951	1.37	.18
(seconds)	RN	2.126	1.729	2.061	1.267	.09	.93
	X	17.818	14.287	21.044	26.501	-.40	.70
	SX	5.137	3.836	3.991	2.419	.73	.47

N number of subjects, *M* mean, *SD* standard deviation, *t* t-value, *p* p-value

Table 5

Comparison between well-tracked and insufficiently tracked ASD participants' RT_{report} in trials 1-5

Group		Well-tracked ASD (*N* = 18)		Insuff. tracked ASD (*N* = 11)			
		M	SD	M	SD	*t*	*p*
RT_{report}	N	4.063	3.977	5.914	4.867	-1.12	.27
(seconds)	RN	2.115	1.650	2.566	0.979	-.82	.42
	X	10.261	11.090	22.261	14.837	-2.19	.02
	SX	4.457	4.314	5.980	3.264	-1.01	.32

ASD Autism Spectrum Disorder, *N* number of subjects, *M* mean, *SD* standard deviation, *t* t-value, *p* p-value

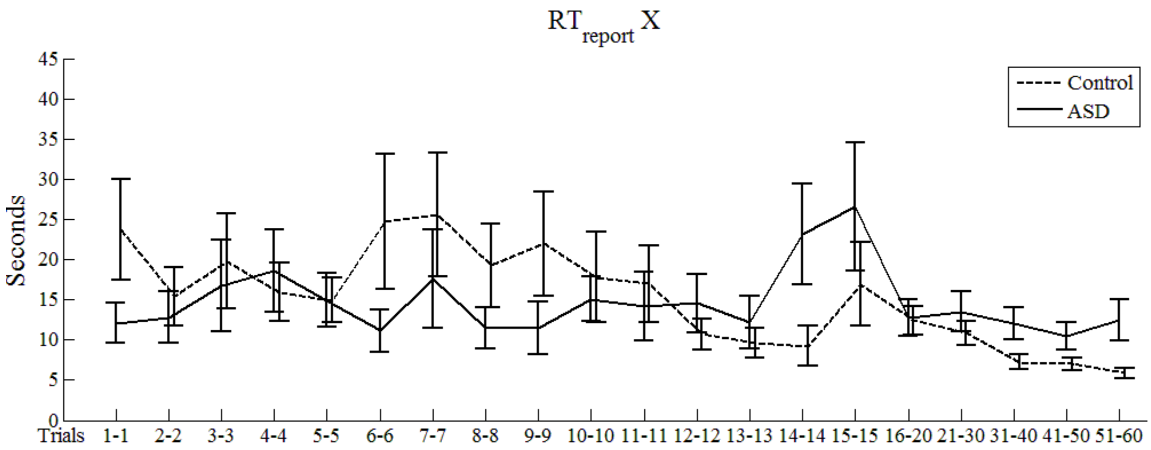


Figure 8. RT$_{report}$ of all 29 ASD and 28 Control subjects regardless of tracking quality. Trial intervals are shown at the x-axis. Error bars show standard error of the mean. Missing values were ignored when averaging over subjects within a group.

Chapter 4 Discussion

4.1 Summary

We sought to compare two processes of visual attention between adults with autism spectrum disorder and a typically developed, matched control group. We used eye-tracking in a visual search for a unique item that differs in its orientation feature from distractor items, but is equal to them on a higher object shape level, giving rise to object-to-feature interference (Zhaoping & Guyader, 2007). Let us restate the assumptions of the measures taken: The experimental paradigm has a measure RT_{eye} that corresponds to visual attentional processes that occur in a bottom-up fashion. This reading of bottom-up refers to those brain processes that involve mostly low-level visual areas, such as V1 (Li, 2002), and that are largely independent of higher level areas. The paradigm yields another measure RT_{lapse} that corresponds to an interference between this bottom-up process and another process, which supposedly proceeds in a top-down manner with respect to brain areas. In particular, it has been suggested that this interference involves viewpoint invariant object recognition: Zhaoping and Guyader (2007) showed that the interference was notable by prolonged RT_{lapse} in X compared with SX-search, and often accompanied by deviation of gaze from the target after first gaze arrival to it. However, even in those trials where gaze stayed on the target until observers reported it, the observers' report was delayed in condition X compared with SX. The authors suggested that a decision process vetoed the first guess by bottom-up feature detectors in X-search because the attended target object was recognized as having the same shape as the distractors. Motor commands can be assumed to be equal across all conditions.

In sum, we did not find a difference between the groups on average over trials with respect to any of the experimental conditions employed. Groups differed, however, with respect to RT_{lapse} in the initial trials of the most difficult search condition X-search, whereas they did not differ in its control condition SX-search. In particular, participants with ASD had lower RT_{lapse} compared to the Control group (while differences in accuracy were insignificant). After the initial trials, the advantage of ASD participants did not disappear because RT_{lapse} increased in the ASD group over time, but instead because the control group quickly managed to reduce this delay (after about 10 trials of this search type). Although this result seems isolated, it is unlikely to have occurred by pure chance. There was a reason to look at initial trials based on a previous study (Oliveri et al., 2010), and we showed that the difference persisted with manipulations of parameters of the data analysis. However, another aspect of the data introduces uncertainty about this finding. It might be the result of a bias in the ASD subjects used for eye-data analysis. These subjects had lower RT_{report} (i.e., the time of button press) compared to 11 ASD subjects excluded from eye-data analysis due to insufficient eye-tracking quality. There were no obvious differences with respect to age, or other participant characteristics, between the well-tracked and the insufficiently tracked ASD subjects. The source of this bias remains unclear; but it could mean that the true ASD population mean of RT_{lapse} might be higher than is estimated by our main analysis. Given these uncertainties, we conclude that there is not enough evidence at this point for a difference in top-down processing between people with autism and those without in our task. Some of the following considerations are therefore tentative.

4.2 Interpretation

The fact that the groups were comparable in their RT_{eye} rules out the hypothesis that bottom-up control of attention works faster in people with autism. Indeed, another visual search study in autism employing eye-tracking did not find an indication of faster bottom-up allocation of attention, which they measured by frequency of fixations, saccade amplitude and distribution of fixations over the search image (Joseph et al., 2009). We decided to test this hypothesis nevertheless, as these authors did not measure the latency of gaze arrival to the visual search target specifically.

What does it mean that there was no significant difference between the groups in the other two search types, N and RN-search, with respect to RT_{lapse}? Behaviorally, the letter searches can be said to be easier, as they consistently yield lower reaction times than X-search. The fact that there was a significant main effect of the condition factor in all our ANOVAs further attests to the difference between search types. Although both N and RN-search do evoke top-down interference they might not be as sensitive a task as X-search in order to reveal differences between an ASD and a control group. Previous studies in autism revealed differences between groups in feature search most reliably when it was difficult or inefficient (e.g. O'Riordan et al., 2001).

What could be a possible explanation of a difference between the groups in RT_{lapse} in X-search? If viewpoint invariance is applied, all X-shaped items in X-search are the same—the target shape does not stand out and can be lost among distractor shapes. However, this top-down process is not required for successful detection of the target—it is even detrimental to it. If this process was weaker to start with or could be somehow resisted (e.g., via other top-down controlling signals), or if, in turn, the bottom-up process was relatively stronger, so that it already determines the perceptual decision, this would reduce the delay in reporting the target location. If autism was associated with weaker top-down processing and/or stronger weight of bottom-up processes on the decision, this would be a possible explanation of such a result.

Based on imaging data from multiple different tasks, C. Frith (2003) speculates that autism features weaker top-down processing (in the brain and in attention): imaging studies show consistently that early sensory processing areas are either normally or over-active, while later processing areas are under-active. He hypothesizes that top-down signals from the attentional system located in parietal and prefrontal cortex could fail to modulate connectivity at lower areas. Findings of weaker expectation or top-down modulated blood oxygen level-dependent response to face stimuli appearing at the place of an exogenous cue for spatial attention may fall in line with this hypothesis (Bird, Catmur, Silani, Frith, & Frith, 2006). The hypothesis of weaker top-down control over the flow of information has yet to be tested in the non-social realm. It should be noted that in the same study modulation of neural response to house stimuli following the exogenous cue was comparable in the two groups, hinting on the possibility that weak top-down control, if a feature of autism at all, might be stimulus-specific.

On the other hand, stronger bottom-up relative to the top-down processes would lead to two outcomes in our task. It has been shown previously that strengthening the bottom-up feature detection process (by making all intersecting bars uniformly horizontal or vertical, see Zhaoping & Guyader, 2007, condition A') reduces both RT_{eye} and RT_{lapse} relative to the RTs of X search (condition A in their study). Thus, if the bottom-up processes are stronger in ASD, we would expect both their RT_{lapse} and, importantly, their RT_{eye} as well, to be lower compared with controls. However, we found no indication of lower RT_{eye} in the ASD group. This means that bottom-up processes do not seem to be relatively stronger in autism, but they could be more independent of top-down influences than in typical people: The Enhanced Perceptual Functioning (EPF) account predicts relative autonomy of perceptual processes from top-down influences, such as of discrimination processes from the top-down influence of categorization (Mottron et al., 2006). In contrast to Frith (2003), EPF does not assume the top-down influences to be weaker per se, as people with autism can successfully solve categorization tasks (Soulières, Mottron, Saumier, & Larochelle, 2007). Nor does it predict stronger bottom-up processes per se, although the model assumes that in general the occipital regions, in which presumably bottom-up processes are at work, are recruited more relative to frontal regions in autism. However, whether this accounts for this autonomy, or what accounts for it, is not a central theme of EPF.

In sum, although bottom-up processes might not be stronger relative to top-down processes in autism, their independency from top-down processes might enable them to make a decision based on bottom-up processes more easily compared with typical observers. It is also possible that top-down processing cannot interfere as much because it is generally weaker in autism (Frith, 2003).

4.3 Possible confounds

4.3.1 Expertise on visual tasks

An anonymous reviewer suggested the possibility that participants with ASD, although coming from the same data base as the control group, might have been exposed to visual tasks more often than the control participants, due to frequent recruitment. This would explain a quicker “tuning in” to the mental schema of participating in a psychological study and perhaps account for immediate acquisition of a strategy to suppress task-irrelevant information. In a post-experimental interview we asked participants for the number of visual studies they had participated in from 2 years prior to participation in ours. Based on these reports, participants with ASD had not been exposed to visual experiments more often than control participants.

4.3.2 Strategic advantage

Observers could be tempted to adopt a strategy of scanning only half of the display in order to make a decision about target location. This strategy works since a target was present in each trial. Thus, scanning only that side of the display which does not contain a target still allows for a correct report, and would supposedly speed up reporting, since fewer items are looked at. Note that this task, although it is a feature search, can be quite inefficient under strong top-down interference, so that search items might be checked in a serial manner. People with ASD could somehow be more inclined to adopt this strategy. However, observers would not gaze at the target in trials where such a strategy was adopted. As we did not analyze trials without gaze arrival to the target for their reaction time, trials where such a strategy was adopted would not weigh into the reaction times analyzed here. The design with only target present trials was chosen, as the alternative of including target absent trials could have devastating effects on the analyzability of data from correct trials: As X-search is very difficult, observers would possibly erroneously discard too many target present trials (Zhaoping, 2014).

4.3.3 Fast learning effects

The control group had learned to resist the top-down process within ten experimental trials. Would it be possible that those participants with ASD who displayed weaker interference had learned the task even faster, i.e., had they learned it even before the start of the experimental trials, based on the four practice trials per search type? Unfortunately we did not record practice trials, which clearly limits the interpretability of the results. However, given the inability of people with ASD to adapt to novel situations (insisting on sameness) and the fact that other types of learning, such as language acquisition, are delayed rather than accelerated in autism, it seems unlikely that a very rapid learning effect after only four exposures to X-search during practice accounts for the low RTs of the ASD group at the beginning of the experimental trials. On the other hand, the learning mechanisms of this task may be completely different from those involved in more complex cognitive abilities, such as language or social behavior. Hence, fast learning must be considered a possible explanation for our finding, but we have to remain agnostic towards its truth. If it were the true explanation, an interesting question for future investigations would be: why do people with ASD achieve these fast learning effects in a visual task, while typical observers do not?

4.3.4 Enhanced processing of detail

All major cognitive theories of autism assert that attention to detail is elevated in ASD compared with the normal population. Claims of higher visual acuity in autism have been refuted (Bach & Dakin, 2009; Falkmer et al., 2011). On the other hand, orientation discrimination as a function of varying luminance contrast has been found to be superior in autism in one study (Bertone et al., 2005), and discrimination threshold of grating orientations with steady contrast was found to decrease as severity of autistic traits in normal subjects increases (Dickinson, Jones, & Milne, 2014). At first sight, it seems advantageous in our task to perceive stimuli with higher acuity or to be able to better discriminate left-and right-slanted bars, as such an ability might make the differently oriented bar “pop out” more. The fact that the time of first gaze arrival to the target was comparable between the groups, however, speaks against such a mechanism. In contrast to this first intuition, a more fine-grained (or higher resolution) perception of the search stimuli might even be a disadvantage: In the original study (Zhaoping & Guyader, 2007) observers developed a strategy of “blurring” their vision to speed up target detection. This strategy works since low spatial frequency (SF) information is processed faster (via the magnocellular, not the slower parvocellular, visual pathway), compared with high SF (detailed) visual input (Frazor, Albrecht, Geisler, & Crane, 2004). There is no evidence, however, that people with autism are more inclined to processing low SF information. In contrast, there have been reports of preferred processing of high SF compared with low SF in autism using faces as stimuli (e.g. Deruelle, Rondan, Gepner, & Tardif, 2004; but see Leonard, Annaz, Karmiloff-Smith, & Johnson, 2011). Two studies found higher sensitivity to high SF (8 cycles per degree) in adults with autistic disorder (Kéïta, Guy, Berthiaume, Mottron, & Bertone, 2014) and higher visual evoked potentials in response to high SF gratings in 4-year olds with ASD (Vlamings, Jonkman, van Daalen, van der Gaag, & Kemner, 2010), respectively, compared with control groups.

4.3.5 Degraded peripheral vision

Could people with autism have a blurred peripheral vision, leading to a similar effect as a “defocusing” strategy, which can be adopted to recruit primarily bottom-up processes to find the target while fixating at the display center? There is preliminary evidence that adolescents with autism show reduced sensitivity in detecting light flashes in the peripheral (30°-85° eccentricity) visual field (Milne, Scope, Griffiths, Codina, & Buckley, 2013). However, this insensitivity is inconsistent with the hypothesis of a gain for bottom-up attention when using peripheral vision. Furthermore, in our experiment, the target appeared within 10° eccentricity from the display center, so that Milne et al.’s results would hardly apply.

4.4 Speculations

4.4.1 Mental rotation in autism

People with autism were tested specifically for object recognition of human bodies and objects (cars) with different viewpoints in a study by Pearson, Marsh, Hamilton and Ropar (2014). Indeed, two studies found superior mental rotation abilities in autism (Falter, Plaisted, & Davies, 2008; Soulières, Zeffiro, Girard, & Mottron, 2011). Faster RTs in the mental rotation task can be interpreted as a faster decision as to whether two items are the same objects or left-right mirror images of each other. In our task, the same decision has to be made when interference leads observers’ gaze away from the target. It is conceivable that fast mental rotation abilities would amount to faster overcoming of confusion between target and distractors in our task, and thereby, could lead to reduced RT_{lapse}. However, as the relationship between the processes that favor performance in the mental rotation task, and the processes involved in our task is not clear, superior mental rotation abilities in autism have to be considered an unrelated finding for the time being.

4.4.2 Object recognition and binding

While object identification and discrimination is “generally reported to be equivalent to those of control populations” and, in fact, is often used as a control task in studies investigating social stimuli (Simmons et al., 2009), only few studies have concentrated on testing top-down processes in the ventral stream in autism. ASD groups demonstrated some difficulty with fine object discrimination in two studies (Behrmann et al., 2006; Davies, Bishop, Manstead, & Tantam, 1994). Two studies report possibly impaired form perception in high-functioning autism but not in Asperger's syndrome (Tsermentseli, O’Brien, & Spencer, 2008), and impaired gestalt grouping in children with autism (Brosnan, Scott, Fox, & Pye, 2004), respectively. However, another study found no differences between a group with high-functioning ASD and a control group with respect to visual functioning at different levels of the ventral stream (de Jonge et al., 2007). Future research is needed to clarify these results.

The idea that people with autism are impaired in the integration of parts of visual objects into wholes has been put forward by some authors. Loth, Gómez, and Happé (2010) found reduced recognition rates of two-tone images of objects and faces after being exposed to the corresponding grey-scale photograph in a group of adults with ASD compared with typical observers. They speculate that this could be an indication of deficits in feature-binding, since priming for two-tone images is thought to rely on top-down processes that increase feature-binding to fill in gaps in impoverished visual input (Dolan et al., 1997; it should be noted that Dolan et al. make a link between two-tone images and feature-binding only implicitly, by showing that the lateral parietal cortex is implicated in recognition of these stimuli, and at the same time happens to be associated with feature-binding). Jolliffe and Baron-Cohen (1997) sought to study the integration of fragments of line drawings for object recognition versus their recognition based on matching only one presented fragment to mental templates in people with autism, Asperger’s syndrome and a normal control group. The ASD groups were able to recognize objects from a single fragment, but were less able to integrate pieces to name the object compared to controls. Olu-Lafe, Liederman, and Tager-Flusberg (2014) found that their ASD group was slower in matching two-part shapes to one of two silhouettes, although they were able to match a single shape to its corresponding silhouette among two alternative silhouettes. These findings suggest that there could be an impaired tendency to integration in people with autism. Note, however, that these three studies tested binding at much higher levels by using far more complex object shapes than those common in visual search. This might explain why difficulties with perceptual integration are not observed in autism with respect to conjunction search—a task which supposedly requires feature integration. The parietal lobe seems to have a role in feature integration (Shafritz, Gore, & Marois, 2002). Two studies on visuospatial attention found reduced parietal lobe activity in subjects with ASD (Belmonte & Yurgelun-Todd, 2003; Haist, Adamo, Westerfield, Courchesne & Townsend, 2005). More studies are needed to investigate parietal cortex activation particularly during visual search in autism. This would be of particular interest since the parietal cortex is involved in inefficient searches (Nobre et al., 2003)—the type of visual search in which differences between autism and typical observers are most robustly found.

4.5 Limitations of the present study

The task design employed here has clear limitations. It cannot distinguish between possible neural sources of the observed delay in X-search relative to SX-search. The top-down interference might originate in the dorsal stream (e.g. at the attentional system in the parietal cortex; spatial attention and/or binding-processes) or in the ventral stream (e.g. V4, IT, where object recognition is usually localized), or both. Oliveri et al. (2010) showed that the right parietal cortex is implicated in this task. Oliveri et al. note that this area seems to be involved in feature integration, as well as in spatial attention, both of which could be important prerequisites for the interference to occur.

Technical issues also limit the interpretability of our results. The fact that we did not obtain good eye data for a substantial part of the participants is another clear limitation of this study. With a particular interest in the earliest trials (including training trials), a design in which the transient effect of strong top-down interference is made to last longer would be desirable. Now that it is clear which search types reveal differences between the groups, multiple testing issues could be avoided by reducing the total number of different conditions.

4.6 Significance of the finding and future direction

Our finding may be taken with caution. We could not preclude the possibility of a speed-accuracy tradeoff nor that of very fast strategic learning effects in the autism group. More importantly, however, we could not explain the finding of a significant difference in the button report times between the well-tracked and the insufficiently tracked ASD participants. Despite these and other limitations, it might be a useful, thought-provoking impulse for other researchers who are interested in the study of the integrity and interaction of bottom-up and top-down object-related visual processing in autism, which is an understudied area at the moment. We alluded to several topics that could be explored further in autism research. Theories have been put forward that predict reduced reliance on top-down processing in autism compared with typical people, either because of autonomy of lower level processing (Mottron et al., 2006) or because feedback processing is thought to be weak (Frith, 2003). Future research should try to test these predictions, and to find out whether they might be stimulus specific or general.

Chapter 5 Conclusion

This thesis set out to achieve two goals: 1) shedding light on results of visual search studies in autism in the literature review in Chapter 2, and 2) answering the question whether people with autism would be facilitated in a visual search task involving a specific interference from higher level processes, and if so, whether this facilitation would be present in bottom-up feature detection or top-down interference.

Chapter 2 reviewed visual search studies in autism published in the past 20 years. The fact that this review is not a systematic review is clearly a limitation of this thesis. Nevertheless, it contributes novel conclusions to the literature which can help point researchers to unresolved questions: Visual search is often said to be superior in autism without making further distinctions between task designs or demographic features. An exception to this is a previously reported pattern of results, which includes that children with autism are robustly faster in conjunction search but not feature search compared with control groups, except when the feature search is difficult enough (O'Riordan et al., 2001). The present review concludes that more recent studies continue to present this pattern of results. In addition, reviewing more recently published studies reveals that findings are mixed with respect to adults and feature search. Conjunction search has to be further investigated in adults with autism, as this type of study is underrepresented in autism research. Differences between groups of subjects with autism and typical groups are most reliably found in inefficient searches—these can be feature as well as conjunction searches. Future investigations should focus on singling out the factors that make the searches inefficient in order to test people with autism specifically with regard to these factors. Note that these conclusions were drawn with respect to visual searches with non-social stimuli (i.e., no faces, people, face drawings etc.), and to searches following a classic design, such as that of Treisman and Gelade (1980).

The experiment described in Chapter 3 sought to test the prediction that people with autism have a superior ability to ignore distracting, task-irrelevant top-down information, compared with typical people, in a visual search task under interference from object recognition. Several findings from a variety of visual tasks inspired this prediction. By utilizing eye-tracking, this study yielded a more exact measure for bottom-up control over attention during visual search than do reaction time recordings using button presses only, and with that, allowed for taking a measure reflecting the top-down processing. We found that people with autism spectrum disorder were comparable to the typical group in terms of bottom-up attentional processes. A finding of possibly reduced top-down interference in ASD compared with controls could not reliably be shown. If future investigations achieve this, such a finding could point towards a general reduction of higher-level object processing in autism, or reduced top-down feedback from later to earlier stages of visual processing. Nevertheless, only future investigations can determine whether this finding is reliable, as the results of the study presented here do not provide an unambiguous picture.

Visual attention remains an interesting topic in autism research, and there are still many unanswered questions. Abnormalities in visual processing, if they are present from an early age on, could influence cognitive development, perhaps accounting for some of the symptomatology of autism (Ames & Fletcher-Watson, 2010). Arguably, there is a major division in autism research between those who favor a “social orienting” theory and those who favor a “bottom-up” theory, focusing on explanations for symptoms of the disorder based on sensory processing abnormalities (Simmons et al., 2009). This distinction translates into a division between “top-down or higher-level deficits involving social cognitive and attentional factors” and “fundamental perceptual mechanisms”, including vision (Greenaway & Plaisted-Grant, 2013). One of the major challenges in the reconciliation of this debate is to identify exactly what is different in visual processing between people with autism and those without.

References

Abrahams, B. S., & Geschwind, D. H. (2008). Advances in autism genetics: on the threshold of a new neurobiology. *Nature Reviews Genetics*, *9*(5), 341-355.

Allport, A. (1993). Attention and control: Have we been asking the wrong questions? A critical review of twenty-five years. *Attention and performance XIV: Synergies in experimental psychology, artificial intelligence, and cognitive neuroscience*, 183-218.

American Psychiatric Association (APA). (2013). *Diagnostic and statistical manual of mental disorders (5th ed., pp. 5–25)*. Arlington, VA: American Psychiatric Publishing.

Ames, C., & Fletcher-Watson, S. (2010). A review of methods in the study of attention in autism. *Developmental Review*, *30*(1), 52-73.

Anderson, B. (2011). There is no such thing as attention. *Frontiers in psychology*, *2*.

Andersson, R., Nyström, M., & Holmqvist, K. (2010). Sampling frequency and eye-tracking measures: how speed affects durations, latencies, and more.*Journal of Eye Movement Research*, *3*(3), 1-12

Bach, M., & Dakin, S. C. (2009). Regarding "Eagle-eyed visual acuity: An experimental investigation of enhanced perception in autism". *Biological Psychiatry*, *66*(10), e19-e20.

Baldassi, S., Pei, F., Megna, N., Recupero, G., Viespoli, M., Igliozzi, R., ... & Cioni, G. (2009). Search superiority in autism within, but not outside the crowding regime. *Vision Research*, *49*(16), 2151-2156.

Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin, J., & Clubley, E. (2001). The autism-spectrum quotient (AQ): Evidence from asperger syndrome/high-functioning autism, malesand females, scientists and mathematicians. *Journal of Autism and Developmental Disorders*, *31*(1), 5-17.

Beauchamp, M. S., Petit, L., Ellmore, T. M., Ingeholm, J., & Haxby, J. V. (2001). A parametric fMRI study of overt and covert shifts of visuospatial attention. *Neuroimage*, *14*(2), 310-321.

Behrmann, M., Avidan, G., Leonard, G. L., Kimchi, R., Luna, B., Humphreys, K., & Minshew, N. (2006). Configural processing in autism and its relationship to face processing. *Neuropsychologia*, *44*(1), 110-129.

Belmonte, M. K., & Yurgelun-Todd, D. A. (2003). Functional anatomy of impaired selective attention and compensatory processing in autism. *Cognitive Brain Research*, *17*(3), 651-664.

Bertone, A., & Faubert, J. (2006). Demonstrations of decreased sensitivity to complex motion information not enough to propose an autism-specific neural etiology. *Journal of Autism and Developmental Disorders*, *36*(1), 55-64.

Bertone, A., Mottron, L., Jelenic, P., & Faubert, J. (2005). Enhanced and diminished visuo-spatial information processing in autism depends on stimulus complexity. *Brain*, *128*(10), 2430-2441.

Bichot, N. P., Rossi, A. F., & Desimone, R. (2005). Parallel and serial neural mechanisms for visual search in macaque area V4. *Science*, *308*(5721), 529-534.

Bichot, N. P., Schall J. D. (2005). Selection and Control of Covert and Overt Orienting. *Neurobiology of Attention.* 117-123.

Bird, G., Catmur, C., Silani, G., Frith, C., & Frith, U. (2006). Attention does not modulate neural responses to social stimuli in autism spectrum disorders.*Neuroimage*, *31*(4), 1614-1624.

Bisley, J. W., & Goldberg, M. E. (2010). Attention, intention, and priority in the parietal lobe. *Annual Review of Neuroscience*, *33*, 1.

Bölte, S., Holtmann, M., Poustka, F., Scheurich, A., & Schmidt, L. (2007). Gestalt perception and local-global processing in high-functioning autism. *Journal of Autism and Developmental Disorders*, *37*(8), 1493-1504.

Brenner, L. A., Turner, K. C., & Müller, R. A. (2007). Eye movement and visual search: are there elementary abnormalities in autism? *Journal of Autism and Developmental Disorders*, *37*(7), 1289-1309.

Brian, J. A., & Bryson, S. E. (1996). Disembedding performance and recognition memory in autism/PDD. *Journal of Child Psychology and Psychiatry*, *37*(7), 865-872.

Broadbent, D. E. (1958). *Perception and communication*. New York: Oxford University Press.

Caron, M. J., Mottron, L., Berthiaume, C., & Dawson, M. (2006). Cognitive mechanisms, specificity and neural underpinnings of visuospatial peaks in autism. *Brain*, *129*(7), 1789-1802.

Chen, L. (1982). Topological structure in visual perception. *Science*,*218*, 699.

Chun, M. M., & Jiang, Y. (1999). Top-down attentional guidance based on implicit learning of visual covariation. *Psychological Science*, *10*(4), 360-365.

Colby, C. L., & Goldberg, M. E. (1999). Space and attention in parietal cortex. *Annual Review of Neuroscience*, *22*(1), 319-349.

Collignon, O., Charbonneau, G., Peters, F., Nassim, M., Lassonde, M., Lepore, F., ... & Bertone, A. (2013). Reduced multisensory facilitation in persons with autism. *Cortex*, *49*(6), 1704-1710.

Constable, P. A., Solomon, J. A., Gaigg, S. B., & Bowler, D. M. (2010). Crowding and visual search in high functioning adults with autism spectrum disorder. *Clinical Optometry*, *2010*(2), 1-11.

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

Corbetta, M., Shulman, G. L., Miezin, F. M., & Petersen, S. E. (1995). Superior parietal cortex activation during spatial attention shifts and visual feature conjunction. *Science*, *270*(5237), 802-805.

Dakin, S., & Frith, U. (2005). Vagaries of visual perception in autism. *Neuron*, *48*(3), 497-507.

Davies, S., Bishop, D., Manstead, A. S., & Tantam, D. (1994). Face perception in children with autism and Asperger's syndrome. *Journal of Child Psychology and Psychiatry*, *35*(6), 1033-1057.

De Jonge, M. V., Kemner, C., De Haan, E. H., Coppens, J. E., Van den Berg, T. J. T. P., & Van Engeland, H. (2007). Visual information processing in high-functioning individuals with autism spectrum disorders and their parents. *Neuropsychology*, *21*(1), 65.

Deruelle, C., Rondan, C., Gepner, B., & Tardif, C. (2004). Spatial frequency and face processing in children with autism and Asperger syndrome. *Journal of Autism and Developmental Disorders*, *34*(2), 199-210.

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

Di Lollo, V., Kawahara, J. I., Zuvic, S. M., & Visser, T. A. (2001). The preattentive emperor has no clothes: A dynamic redressing. *Journal of Experimental Psychology: General*, *130*(3), 479.

Dickinson, A., Jones, M., & Milne, E. (2014). Oblique orientation discrimination thresholds are superior in those with a high level of autistic traits. *Journal of Autism and Developmental Disorders*, *44*(11), 2844-50.

Dienes, Z. (2014). Using Bayes to get the most out of non-significant results. *Frontiers in Psychology*, *5*, 781.

Dolan, R. J., Fink, G. R., Rolls, E., Booth, M., Holmes, A., Frackowiak, R. S. J., & Friston, K. (1997). How the brain learns to see objects and faces in an impoverished context. *Nature*, *389*(6651), 596-599.

Duncan, J., & Humphreys, G. W. (1989). Visual search and stimulus similarity. *Psychological Review*, *96*(3), 433.

Egeth, H. E., Virzi, R. A., & Garbart, H. (1984). Searching for conjunctively defined targets. *Journal of Experimental Psychology: Human Perception and Performance*, *10*(1), 32.

Falkmer, M., Stuart, G. W., Danielsson, H., Bram, S., Lönebrink, M., & Falkmer, T. (2011). Visual acuity in adults with asperger’s syndrome: No evidence for “eagle-eyed” vision. *Biological Psychiatry*, *70*(9), 812-816.

Falter, C. M., Plaisted, K. C., & Davis, G. (2008). Visuo-spatial processing in autism—Testing the predictions of extreme male brain theory. *Journal of Autism and Developmental Disorders*, *38*(3), 507-515.

Fecteau, J. H., & Munoz, D. P. (2006). Salience, relevance, and firing: a priority map for target selection. *Trends in Cognitive Sciences*, *10*(8), 382-390.

Felleman, D. J., & Van Essen, D. C. (1991). Distributed hierarchical processing in the primate cerebral cortex. *Cerebral Cortex*, *1*(1), 1-47.

Findlay, J. M., & Gilchrist, I. D. (1998). Eye Guidance and Visual Search. In G. Underwood (Ed.), *Eye Guidance in Reading and Scene Perception*, p. 295-312, Oxford, England: Elsevier.

Fockert, J., Rees, G., Frith, C., & Lavie, N. (2004). Neural correlates of attentional capture in visual search. *Cognitive Neuroscience, Journal of*, *16*(5), 751-759.

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.

Frazor, R. A., Albrecht, D. G., Geisler, W. S., & Crane, A. M. (2004). Visual cortex neurons of monkeys and cats: temporal dynamics of the spatial frequency response function. *Journal of Neurophysiology*, *91*(6), 2607-2627.

Frith, C. (2003). What do imaging studies tell us about the neural basis of autism. *Autism: Neural Basis and Treatment Possibilities*, 149-176.

Frith, C. (2005). The top in top-down attention. *Neurobiology of attention*, 105-108

Frith, U. (1974). Acurious effect with reversed letters explained by a theory of schema. *Perception & Psychophysics*, *16*(1), 113-116.

Frith, U. (1989). *Autism: Explaining the enigma.* Blackwell Publishing.

Frith, U., & Happé, F. (1994). Autism: Beyond “theory of mind”. *Cognition*, *50*(1), 115-132.

Goodale, M. A., & Milner, A. D. (1992). Separate visual pathways for perception and action. *Trends in Neurosciences*, *15*(1), 20-25.

Greenaway, R., Davis, G., & Plaisted-Grant, K. (2013). Marked selective impairment in autism on an index of magnocellular function. *Neuropsychologia, 51,* 592-600.

Grinter, E. J., Maybery, M. T., & Badcock, D. R. (2010). Vision in developmental disorders: Is there a dorsal stream deficit?. *Brain Research Bulletin*, *82*(3), 147-160.

Happé, F. G. (1996). Studying weak central coherence at low levels: Children with autism do not succumb to visual illusions. A research note. *Journal of Child Psychology and Psychiatry*, *37*(7), 873-877.

Happé, F. (1999). Autism: cognitive deficit or cognitive style? *Trends in Cognitive Sciences*, *3*(6), 216-222.

Happé, F. G., & Booth, R. D. (2008). The power of the positive: Revisiting weak coherence in autism spectrum disorders. *The Quarterly Journal of Experimental Psychology*, *61*(1), 50-63.

Happé, F., & Frith, U. (2006). The weak coherence account: Detail-focused cognitive style in autism spectrum disorders. *Journal of Autism and Developmental Disorders*, *36*(1), 5-25.

Henderson, J. M. (2003). Human gaze control during real-world scene perception. *Trends in Cognitive Sciences*, *7*(11), 498-504.

Hessels, R. S., Hooge, I. T., Snijders, T. M., & Kemner, C. (2014). Is there a limit to the superiority of individuals with ASD in visual search? *Journal of Autism and Developmental Disorders*, *44*(2), 443-451.

Hochstein, S., & Ahissar, M. (2002). View from the top: Hierarchies and reverse hierarchies in the visual system. *Neuron*, *36*(5), 791-804.

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

Hutton, S. B. (2008). Cognitive control of saccadic eye movements. *Brain and Cognition*, *68*(3), 327-340.

Iarocci, G., Burack, J. A., Shore, D. I., Mottron, L., & Enns, J. T. (2006). Global–local visual processing in high functioning children with autism: Structural vs. implicit task biases. *Journal of Autism and Developmental Disorders*, *36*(1), 117-129.

Itti, L. (2005). Models of bottom-up attention and saliency. *Neurobiology of attention*, 576-582.

Itti, L., & Koch, C. (2000). A saliency-based search mechanism for overt and covert shifts of visual attention. *Vision Research*, *40*(10), 1489-1506.

Itti, L., & Koch, C. (2001). Computational modelling of visual attention. *Nature Reviews Neuroscience*, *2*(3), 194-203.

Itti, L., Koch, C., & Niebur, E. (1998). A model of saliency-based visual attention for rapid scene analysis. *IEEE Transactions on pattern analysis and machine intelligence*, *20*(11), 1254-1259.

Itti, L., Rees, G., Tsotsos, J. K., (2005). Preface, *Neurobiology of attention*.

Jarrold, C., Gilchrist, I. D., & Bender, A. (2005). Embedded figures detection in autism and typical development: Preliminary evidence of a double dissociation in relationships with visual search. *Developmental Science*, *8*(4), 344-351.

Jolliffe, T., & Baron-Cohen, S. (1997). Are people with autism and Asperger syndrome faster than normal on the Embedded Figures Test? *Journal of Child Psychology and Psychiatry*, *38*(5), 527-534.

Jones, C. R., Swettenham, J., Charman, T., Marsden, A. J., Tregay, J., Baird, G., ... & Happé, F. (2011). No evidence for a fundamental visual motion processing deficit in adolescents with autism spectrum disorders. *Autism Research*, *4*(5), 347-357.

Joseph, J. S., Chun, M. M., & Nakayama, K. (1997). Attentional requirements in a 'preattentive' feature search task. *Nature*, *387*(6635), 805-807.

Joseph, R. M., Keehn, B., Connolly, C., Wolfe, J. M., & Horowitz, T. S. (2009). Why is visual search superior in autism spectrum disorder? *Developmental Science, 12*(6), 1083-1096.

Julesz, B. (1984). A Brief Outline of the Texton Theory of Human-Vision. *Trends in Neurosciences, 7*, 41-45.

Kahneman, D., & Beatty, J. (1966). Pupil diameter and load on memory.*Science*, *154*(3756), 1583-1585.

Kaldy, Z., Giserman, I., Carter, A. S., & Blaser, E. (2013). The mechanisms underlying the ASD advantage in visual search. *Journal of Autism and Developmental Disorders*, 1-15.

Kaldy, Z., Kraper, C., Carter, A. S., & Blaser, E. (2011). Toddlers with autism spectrum disorder are more successful at visual search than typically developing toddlers. *Developmental Science, 14*(5), 980-988.

Kanner, L. (1951). The conception of wholes and parts in early infantile autism. *The American Journal of Psychiatry, 108*(1), 23–26.

Kaplan, M., Rimland, B., & Edelson, S. M. (1999). Strabismus in autism spectrum disorder. *Focus on Autism and Other Developmental Disabilities*, *14*(2), 101-105.

Keehn, B., & Joseph, R. M. (2008). Impaired prioritization of novel onset stimuli in autism spectrum disorder. *Journal of Child Psychology and Psychiatry*, *49*(12), 1296-1303.

Keehn, B., Shih, P., Brenner, L. A., Townsend, J., & Müller, R. A. (2013). Functional connectivity for an “Island of sparing” in autism spectrum disorder: an fMRI study of visual search. *Human Brain Mapping*, *34*(10), 2524-2537.

Kéïta, L., Guy, J., Berthiaume, C., Mottron, L., & Bertone, A. (2014). An early origin for detailed perception in Autism Spectrum Disorder: biased sensitivity for high-spatial frequency information. *Scientific reports*, *4*, 5475.

Kemner, C., Van Ewijk, L., Van Engeland, H., & Hooge, I. (2008). Brief report: Eye movements during visual search tasks indicate enhanced stimulus discriminability in subjects with PDD. *Journal of Autism and Developmental Disorders*, *38*(3), 553-557.

Koch, C., & Ullman, S. (1985). Shifts in selective visual attention: towards the underlying neural circuitry. *Human Neurobiology, 4* (1985), pp. 219–227

Kowler, E., Anderson, E., Dosher, B., & Blaser, E. (1995). The role of attention in the programming of saccades. *Vision Research*, *35*(13), 1897-1916.

Leonard, H. C., Annaz, D., Karmiloff-Smith, A., & Johnson, M. H. (2011). Brief report: Developing spatial frequency biases for face recognition in autism and Williams syndrome. *Journal of Autism and Developmental Disorders*, *41*(7), 968-973.

Li, Z. (2002). A saliency map in primary visual cortex. *Trends in Cognitive Sciences*, *6*(1), 9-16.

Liu, Y., Cherkassky, V. L., Minshew, N. J., & Just, M. A. (2011). Autonomy of lower-level perception from global processing in autism: evidence from brain activation and functional connectivity. *Neuropsychologia*, *49*(7), 2105-2111.

Livingstone, M., & Hubel, D. (1988). Segregation of form, color, movement, and depth: anatomy, physiology, and perception. *Science*, *240*(4853), 740-749.

Lord, C., Risi, S., Lambrecht, L., Cook Jr, E. H., Leventhal, B. L., DiLavore, P. C., ... & Rutter, M. (2000). The Autism Diagnostic Observation Schedule—Generic: A standard measure of social and communication deficits associated with the spectrum of autism. *Journal of Autism and Developmental Disorders,30*(3), 205-223.

Loth, E., Gómez, J. C., & Happé, F. (2010). When seeing depends on knowing: adults with autism spectrum conditions show diminished top-down processes in the visual perception of degraded faces but not degraded objects. *Neuropsychologia*, *48*(5), 1227-1236.

Maljkovic, V., & Nakayama, K. (1994). Priming of pop-out: I. Role of features. *Memory & Cognition*, *22*(6), 657-672.

Maljkovic, V., & Nakayama, K. (2000). Priming of popout: III. A short-term implicit memory system beneficial for rapid target selection. *Visual Cognition*,*7*(5), 571-595.

McConkie, G. W., & Currie, C. B. (1996). Visual stability across saccades while viewing complex pictures. Journal of Experimental Psychology: Human Perception & Performance, 22(3), 563-581.

McPeek, R. M., & Keller, E. L. (2004). Deficits in saccade target selection after inactivation of superior colliculus. *Nature Neuroscience*, *7*(7), 757-763.

Melloni, L., van Leeuwen, S., Alink, A., & Müller, N. G. (2012). Interaction between bottom-up saliency and top-down control: how saliency maps are created in the human brain. *Cerebral Cortex*, *22*(12), 2943-2952.

Milne, E., Griffiths, H., Buckley, D., & Scope, A. (2009). Vision in children and adolescents with autistic spectrum disorder: Evidence for reduced convergence. *Journal of Autism and Developmental Disorders*, *39*(7), 965-975.

Milne, E., Scope, A., Griffiths, H., Codina, C., & Buckley, D. (2013). Brief Report: Preliminary Evidence of Reduced Sensitivity in the Peripheral Visual Field of Adolescents with Autistic Spectrum Disorder. *Journal of Autism and Developmental Disorders*, *43*(8), 1976-1982.

Milne, E., Swettenham, J., Hansen, P., Campbell, R., Jeffries, H., & Plaisted, K. (2002). High motion coherence thresholds in children with autism. *Journal of Child Psychology and Psychiatry*, *43*(2), 255-263.

Milne, E., & Szczerbinski, M. (2009). Global and local perceptual style, field-independence, and central coherence: An attempt at concept validation. *Advances in Cognitive psychology*, *5*, 1-26.

Milne, E., White, S., Campbell, R., Swettenham, J., Hansen, P., & Ramus, F. (2006). Motion and form coherence detection in autistic spectrum disorder: Relationship to motor control and 2: 4 digit ratio. *Journal of Autism and Developmental Disorders*, *36*(2), 225-237.

Mishkin, M., Ungerleider, L. G., & Macko, K. A. (1983). Object vision and spatial vision: Two cortical pathways. *Trends in Neurosciences*, *6*, 414-417.

Mitchell, P., Mottron, L., Soulières, I., & Ropar, D. (2010). Susceptibility to the Shepard illusion in participants with autism: reduced top-down influences within perception? *Autism Research*, *3*(3), 113-119.

Mitchell, P., & Ropar, D. (2004). Visuo-spatial abilities in autism: A review. *Infant and Child Development*, *13*(3), 185-198.

Mottron, L., Belleville, S., & Ménard, E. (1999). Local bias in autistic subjects as evidenced by graphic tasks: Perceptual hierarchization or working memory deficit? *Journal of Child Psychology and Psychiatry*, *40*(05), 743-755.

Mottron, L., Bouvet, L., Bonnel, A., Samson, F., Burack, J. A., Dawson, M., & Heaton, P. (2013). Veridical mapping in the development of exceptional autistic abilities. *Neuroscience & Biobehavioral Reviews*, *37*(2), 209-228.

Mottron, L., & Burack, J. A. (2001). Enhanced perceptual functioning in the development of autism. *The development of autism: Perspectives from theory and research*, 131-148.

Mottron, L., Burack, J. A., Iarocci, G., Belleville, S., & Enns, J. T. (2003). Locally oriented perception with intact global processing among adolescents with high-functioning autism: Evidence from multiple paradigms. *Journal of Child Psychology and Psychiatry*, *44*(6), 904-913.

Mottron, L., Dawson, M., Soulieres, I., Hubert, B., & Burack, J. (2006). Enhanced perceptual functioning in autism: an update, and eight principles of autistic perception. *Journal of Autism and Developmental Disorders*, *36*(1), 27-43.

Mottron, L., & Soulières, I. (2013). Global Versus Local Processing.*Encyclopedia of Autism Spectrum Disorders*, 1445-1451.

Müller, H. J., & Krummenacher, J. (2006). Visual search and selective attention. *Visual Cognition*, *14*(4-8), 389-410.

Nakayama, K., Joseph, J. S., & Parasuraman, R. (1998). Attention, pattern recognition and popout in visual search. *The attentive brain*.

Nakayama, K., & Martini, P. (2011). Situating visual search. *Vision Research*, *51*(13), 1526-1537.

Neisser, U. (1967). *Cognitive psychology*. New York: Appleton-Century-Crofts.

Newschaffer, C. J., Croen, L. A., Daniels, J., Giarelli, E., Grether, J. K., Levy, S. E., ... & Windham, G. C. (2007). The epidemiology of autism spectrum disorders*. *Annual Review of Public Health*, *28*, 235-258.

Nobre, A. C., Coull, J. T., Walsh, V., & Frith, C. D. (2003). Brain activations during visual search: Contributions of search efficiency versus feature binding. *Neuroimage*, *18*(1), 91-103.

Nobre, A. C., Gitelman, D. R., Dias, E. C., & Mesulam, M. M. (2000). Covert visual spatial orienting and saccades: overlapping neural systems. *Neuroimage*, *11*(3), 210-216.

Nothdurft, H. C. (2000). Salience from feature contrast: Additivity across dimensions. *Vision Research*, *40*(10), 1183-1201.

Nothdurft, H. C. (2005). Salience of feature contrast. *Neurobiology of Attention.* 233-239.

Ogawa, T., & Komatsu, H. (2004). Target selection in area V4 during a multidimensional visual search task. *The Journal of Neuroscience*, *24*(28), 6371-6382.

Olu-Lafe, O., Liederman, J., & Tager-Flusberg, H. (2014). Is the ability to integrate parts into wholes affected in Autism Spectrum Disorder? *Journal of Autism and Developmental Disorders*, 1-9.

O'Riordan, M. (2000). Superior modulation of activation levels of stimulus representations does not underlie superior discrimination in autism. *Cognition*,*77*(2), 81-96.

O'Riordan, M. A. (2004). Superior visual search in adults with autism. *Autism*, *8*(3), 229-248.

O'Riordan, M., & Plaisted, K. (2001). Enhanced discrimination in autism. *The Quarterly Journal of Experimental Psychology: Section A*, *54*(4), 961-979.

O'Riordan, M. A., Plaisted, K. C., Driver, J., & Baron-Cohen, S. (2001). Superior visual search in autism. *Journal of Experimental Psychology: Human Perception and Performance*, *27*(3), 719.

Pearson, A., Marsh, L., Hamilton, A., & Ropar, D. (2014). Spatial Transformations of Bodies and Objects in Adults with Autism Spectrum Disorder. *Journal of autism and developmental disorders*, *44*(9), 2277-2289.

Pellicano, E., Gibson, L., Maybery, M., Durkin, K., & Badcock, D. R. (2005). Abnormal global processing along the dorsal visual pathway in autism: A possible mechanism for weak visuospatial coherence? *Neuropsychologia*, *43*(7), 1044-1053.

Peterson, M. S., Kramer, A. F., & Irwin, D. E. (2004). Covert shifts of attention precede involuntary eye movements. *Perception & psychophysics*, *66*(3), 398-405.

Plaisted, K. C. (2001). Reduced generalization in autism: An alternative to weak central coherence. *The development of autism: Perspectives from theory and research,* 149-169.

Plaisted, K., O'Riordan, M., & Baron-Cohen, S. (1998). Enhanced visual search for a conjunctive target in autism: A research note. *Journal of Child Psychology and Psychiatry*, *39*(05), 777-783.

Plaisted, K., Saksida, L., Alcántara, J., & Weisblatt, E. (2003). Towards an understanding of the mechanisms of weak central coherence effects: Experiments in visual configural learning and auditory perception. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, *358*(1430), 375-386.

Quinlan, P. T. (2003). Visual feature integration theory: past, present, and future. *Psychological Bulletin*, *129*(5), 643.

Remington, A., Swettenham, J., Campbell, R., & Coleman, M. (2009). Selective attention and perceptual load in autism spectrum disorder. *Psychological Science, 20*(11), 1388-1393.

Remington, A. M., Swettenham, J. G., & Lavie, N. (2012). Lightening the load: perceptual load impairs visual detection in typical adults but not in autism. *Journal of abnormal psychology*, *121*(2), 544.

Ropar, D., & Mitchell, P. (1999). Are individuals with autism and Asperger's syndrome susceptible to visual illusions? *Journal of Child Psychology and Psychiatry*, *40*(8), 1283-1293.

Ropar, D., & Mitchell, P. (2001). Susceptibility to illusions and performance on visuospatial tasks in individuals with autism. *Journal of Child Psychology and Psychiatry*, *42*(4), 539-549.

Ropar, D., & Mitchell, P. (2002). Shape constancy in autism: The role of prior knowledge and perspective cues. *Journal of Child Psychology and Psychiatry*, *43*(5), 647-653.

Sagi, D. & Julesz, B. (1985). Where and what in vision. *Science*, *228*, 1217-1219.

Sagi, D., & Julesz, B. (1987). Short-range limitation on detection of feature differences. *Spatial Vision*, *2*(1), 39-49.

Scharre, J. E., & Creedon, M. P. (1992). Assessment of visual function in autistic children. *Optometry & Vision Science*, *69*(6), 433-439.

Schmider, E., Ziegler, M., Danay, E., Beyer, L., & Bühner, M. (2010). Is it really robust? Reinvestigating the robustness of ANOVA against violations of the normal distribution assumption. *Methodology: European Journal of Research Methods for the Behavioral and Social Sciences*, *6*(4), 147.

Schmolesky, M. T., Wang, Y., Hanes, D. P., Thompson, K. G., Leutgeb, S., Schall, J. D., & Leventhal, A. G. (1998). Signal timing across the macaque visual system. *Journal of Neurophysiology*, *79*(6), 3272-3278.

Serences, J. T., & Yantis, S. (2006). Selective visual attention and perceptual coherence. *Trends in Cognitive Sciences*, *10*(1), 38-45.

Shafritz, K. M., Gore, J. C., & Marois, R. (2002). The role of the parietal cortex in visual feature binding. *Proceedings of the National Academy of Sciences*,*99*(16), 10917-10922.

Shah, A., & Frith, U. (1983). An islet of ability in autistic children: A research note. *Journal of Child Psychology and Psychiatry*, *24*(4), 613-620.

Shah, A., & Frith, U. (1993). Why do autistic individuals show superior performance on the block design task? *Journal of Child Psychology and Psychiatry*, *34*(8), 1351-1364.

Shipp, S. (2004). The brain circuitry of attention. *Trends in Cognitive Sciences*,*8*(5), 223-230.

Simmons, D. R., Robertson, A. E., McKay, L. S., Toal, E., McAleer, P., & Pollick, F. E. (2009). Vision in autism spectrum disorders. *Vision Research*,*49*(22), 2705-2739.

Smilek, D., Enns, J. T., Eastwood, J. D., & Merikle, P. M. (2006). Relax! Cognitive strategy influences visual search. *Visual Cognition*, *14*(4-8), 543-564.

Soulières, I., Mottron, L., Saumier, D., & Larochelle, S. (2007). Atypical categorical perception in autism: Autonomy of discrimination? *Journal of Autism and Developmental Disorders*, *37*(3), 481-490.

Soulieres, I., Zeffiro, T. A., Girard, M. L., & Mottron, L. (2011). Enhanced mental image mapping in autism. *Neuropsychologia*, *49*(5), 848-857.

Spencer, J., O'Brien, J., Riggs, K., Braddick, O., Atkinson, J., & Wattam-Bell, J. (2000). Motion processing in autism: Evidence for a dorsal stream deficiency. *Neuroreport*, *11*(12), 2765-2767.

Taylor, S., & Badcock, D. (1988). Processing feature density in preattentive perception. *Perception & Psychophysics*, *44*(6), 551-562.

Theeuwes, J. (1991). Exogenous and endogenous control of attention: The effect of visual onsets and offsets. *Perception & Psychophysics*, *49*(1), 83-90.

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.

Treisman, A. (1985). Preattentive processing in vision. *Computer Vision, Graphics, and Image Processing*, *31*(2), 156-177.

Treisman, A. (1988). Features and objects: The fourteenth Bartlett memorial lecture. *The quarterly Journal of Experimental Psychology*, *40*(2), 201-237.

Treisman, A. M. & Gelade, G. (1980). Feature-Integration Theory of Attention. *Cognitive Psychology*, 12, 97-136.

Treue, S., & Cesar Martinez-Trujillo, J. (2006). Visual search and single-cell electrophysiology of attention: Area MT, from sensation to perception. *Visual Cognition*, *14*(4-8), 898-910.

Tribull, M. (2011). *Are people with high-functioning autism or autism spectrum disorder less vulnerable to the reversed letter effect?* (Unpublished bachelor’s thesis). University of Osnabrück, Germany.

Tsermentseli, S., O’Brien, J. M., & Spencer, J. V. (2008). Comparison of form and motion coherence processing in autistic spectrum disorders and dyslexia. *Journal of Autism and Developmental Disorders*, *38*(7), 1201-1210.

Ullman, S. (1984). Visual routines. *Cognition*, *18*(1), 97-159.

Ungerleider, L. G., & Haxby, J. V. (1994). ‘What’and ‘where’ in the human brain. *Current Opinion in Neurobiology*, *4*(2), 157-165.

Vandenbroucke, M. W., Scholte, H. S., van Engeland, H., Lamme, V. A., & Kemner, C. (2008). A neural substrate for atypical low-level visual processing in autism spectrum disorder. *Brain*, *131*(4), 1013-1024.

Vlamings, P. H. J. M., Jonkman, L. M., van Daalen, E., van der Gaag, R. J., & Kemner, C. (2010). Basic abnormalities in visual processing affect face processing at an early age in autism spectrum disorder. *Biological Psychiatry*, *68*(12), 1107-1113.

Wang, Q., Cavanagh, P., & Green, M. (1994). Familiarity and pop-out in visual search. *Perception & Psychophysics*, *56*(5), 495-500.

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

Wolfe, J. M. (1996). Extending guided search: Why guided search needs a preattentive “item map”. *Converging operations in the study of visual selective attention*, 247-270.

Wolfe, J. M. (1998). Visual search. *Attention*, 1, 13-73.

Wolfe, J. M. (2005). Guidance of visual search by preattentive information. *Neurobiology of attention*, 101-104.

Wolfe, J. M. (2003). Moving towards solutions to some enduring controversies in visual search. *Trends in Cognitive Sciences*, *7*(2), 70-76.

Wolfe, J. M. (2014). Approaches to Visual Search: Feature Integration Theory and Guided Search. *The Oxford Handbook of Attention*, 11.

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.

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.

Wolfe, J. M., Horowitz, T. S., Kenner, N., Hyle, M., & Vasan, N. (2004). How fast can you change your mind? The speed of top-down guidance in visual search. *Vision Research*, *44*(12), 1411-1426.

Wolfe, J. M., Treisman, A., & Horowitz, T. S. (2003). What shall we do with the preattentive processing stage: Use it or lose it. *Journal of Vision*, *3*(9), 572.

Wolfe, J. M., Võ, M. L. H., Evans, K. K., & Greene, M. R. (2011). Visual search in scenes involves selective and nonselective pathways. *Trends in Cognitive Sciences*, *15*(2), 77-84.

Wurtz, R. H., & Kandel, E. R. (2000). Constructing the visual image. *Principles of neural science*, 492-506.

Yarbus, A. L. (1967). Eye movements during perception of complex objects. In*Eye movements and vision* (pp. 171-211). Springer US.

Zeki, S. M. (1978). Functional specialisation in the visual cortex of the rhesus monkey. *Nature*,*274*(5670), 423-428.

Zhang, X., Zhaoping, L., Zhou, T., & Fang, F. (2012). Neural activities in V1 create a bottom-up saliency map. *Neuron*, *73*(1), 183-192.

Zhaoping, L. (2005). The primary visual cortex creates a bottom-up saliency map. *Neurobiology of attention.* 570-575.

Zhaoping, L. (2014). *Understanding vision: theory, models, and data*. Oxford University Press.

Zhaoping, L., & Frith, U. (2011). A clash of bottom-up and top-down processes in visual search: The reversed letter effect revisited. *Journal of Experimental Psychology: Human Perception and Performance*, *37*(4), 997.

Zhaoping, L., & Guyader, N. (2007). Interference with bottom-up feature detection by higher-level object recognition. *Current Biology*, *17*(1), 26-31.

Appendices

Appendix A: Example of gaze traces

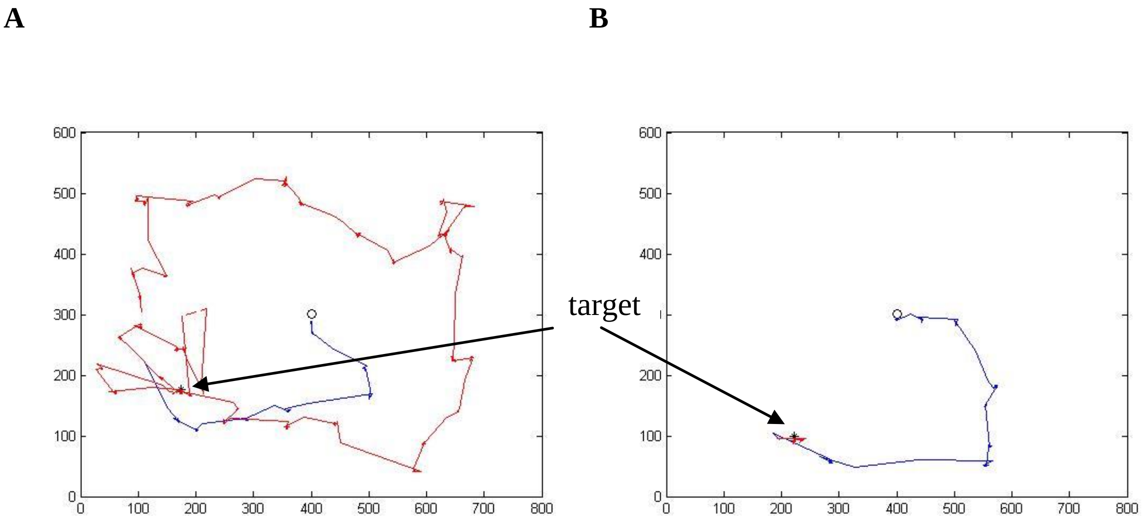


Figure A1. An observer’s gaze traces during X-search (A) and SX-search (B). The blue traces are eye positions before gaze arrived at the target for the first time in the trial. Red traces are eye postions after arrival at the target. Numbers on the panel frames are screen coordinates in pixels. The fixation point is plotted here for reference as a circle in the panel’s center, but was not present in the stimulus.

Appendix B: Characteristics of insufficiently tracked participants

Table B1 *Characteristics of insufficiently tracked participants (N = 18)*

	ASD (*N* = 11) Mean (SD)		Control (*N* = 7) Mean (SD)	
Female:male	2:9		5:2	
Age (years)	40	(11)	38	(6)
vIQ	116.64	(9.95)	110.00	(6.06)
pIQ	107.00	(14.72)	106.29	(13.20)
AQ *	38.91	(7.83)	19.14	(3.02)
ADOS score	9.90	(2.98)		-
BD score	12.18	(3.57)	11.57	(3.46)

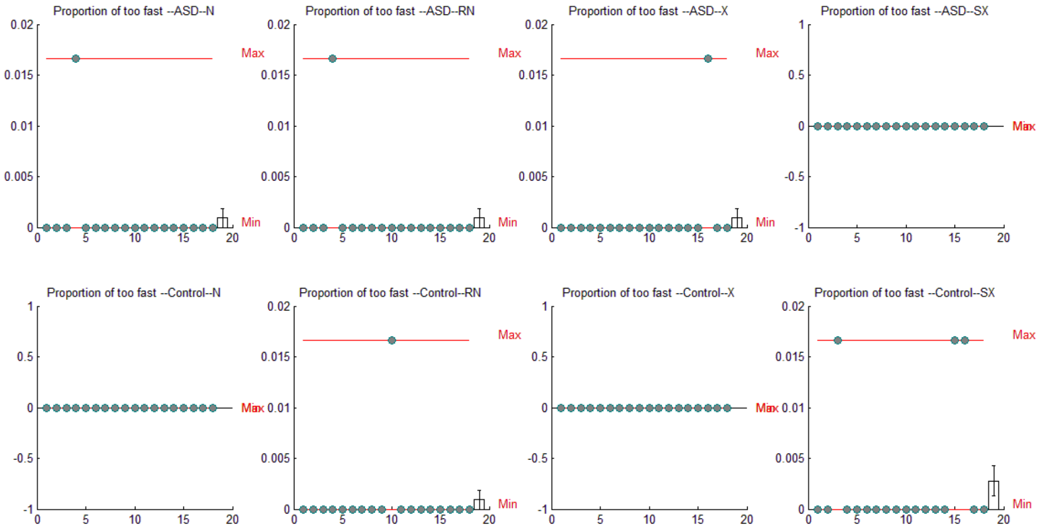
SD standard deviation, *ASD* Autism Spectrum Disorder, *N* number of participants, *vIQ* verbal intelligence quotient, *pIQ* performance intelligence quotient, *AQ* autism quotient, *ADOS* autism diagnostic observation schedule, *BD* Block Design

*$p < .001$ by unpaired, two-tailed t-test

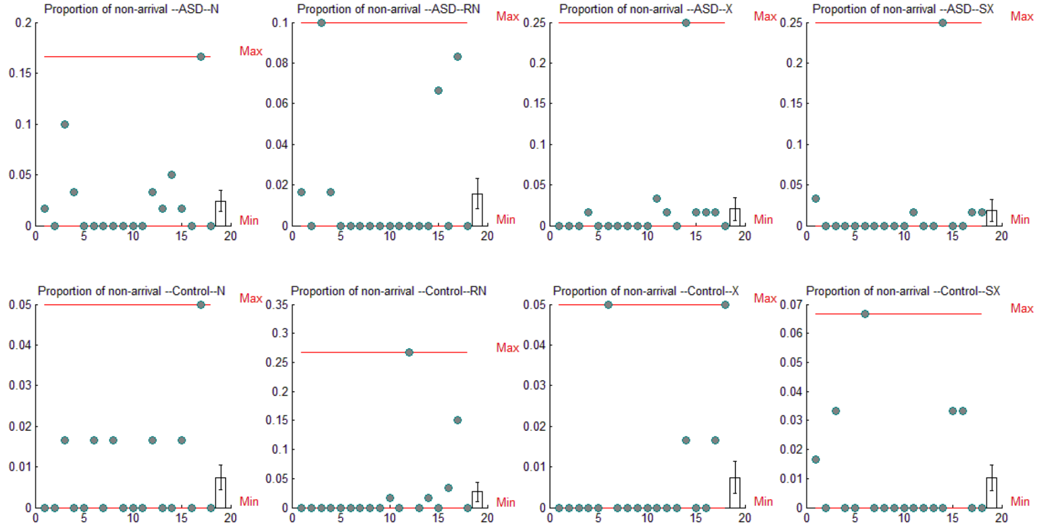
Appendix C: Proportions of excluded and of valid trials

Proportions of excluded and of valid trials among all trials—individual and means

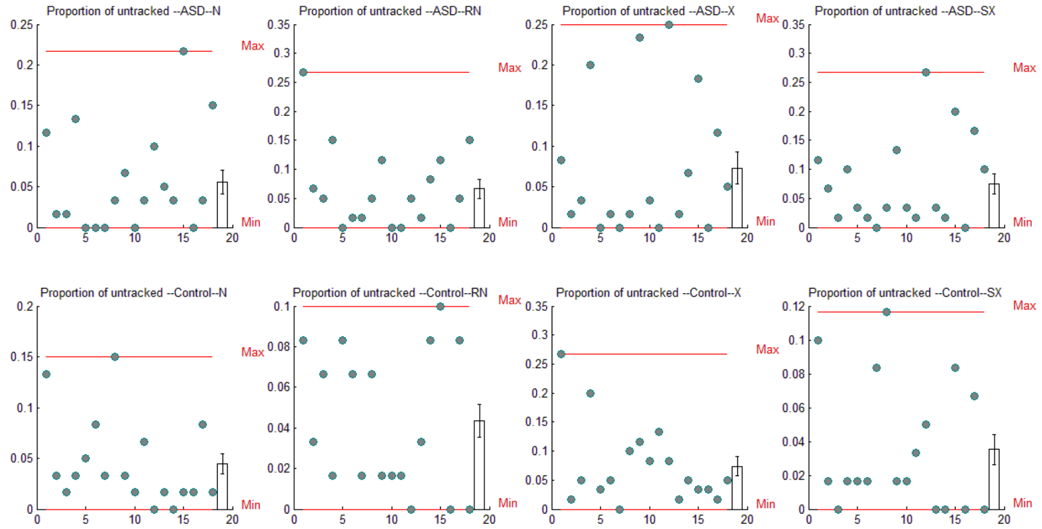
A



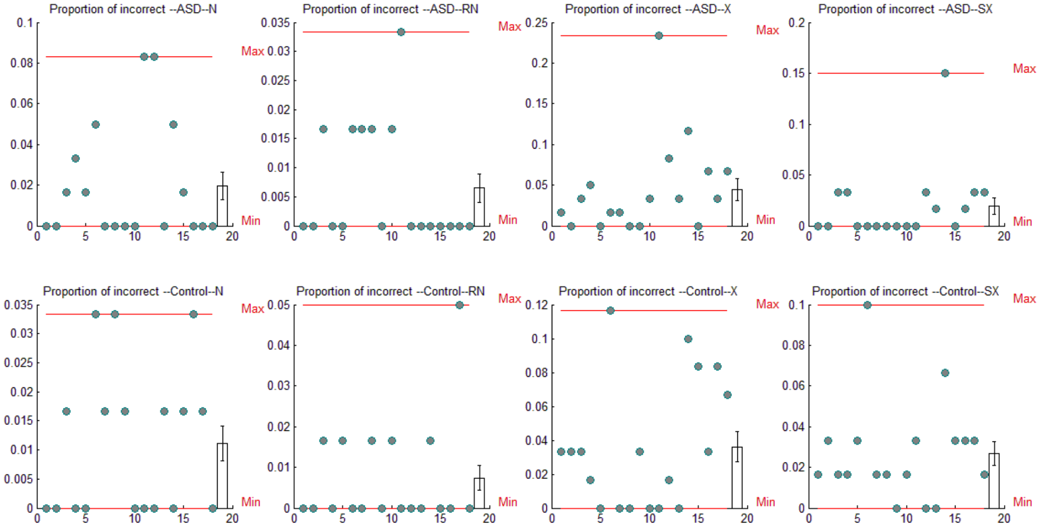
B



C


Proportion of untracked --ASD--N
Proportion of untracked --ASD--RN
Proportion of untracked --ASD--X
Proportion of untracked --ASD--SX
Proportion of untracked --Control--N
Proportion of untracked --Control--RN
Proportion of untracked --Control--X
Proportion of untracked --Control--SX
Max
Min
0
5
10
15
20

D


Proportion of incorrect --ASD--N
Proportion of incorrect --ASD--RN
Proportion of incorrect --ASD--X
Proportion of incorrect --ASD--SX
Proportion of incorrect --Control--N
Proportion of incorrect --Control--RN
Proportion of incorrect --Control--X
Proportion of incorrect --Control--SX
Max
Min
0
5
10
15
20

E

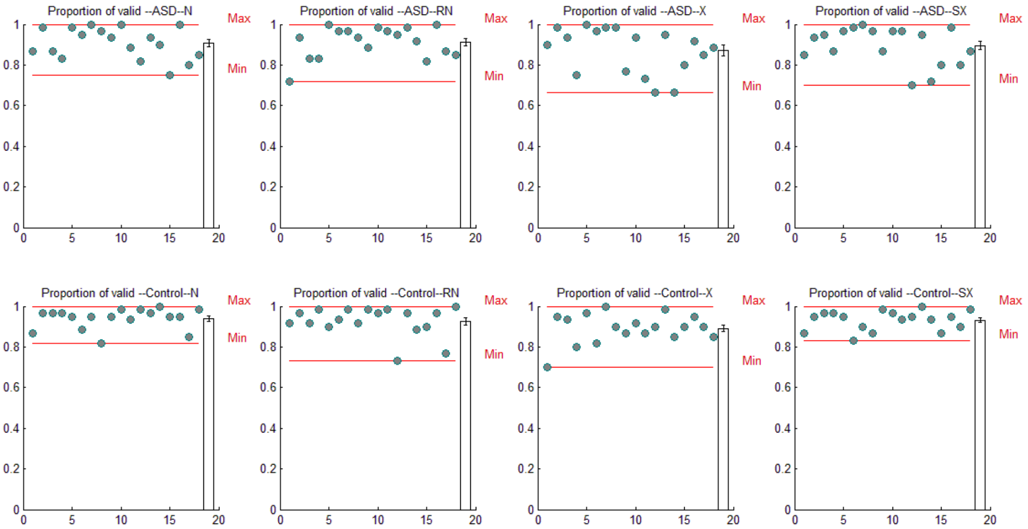


Figure C1. Proportions of excluded and of valid trials. A-E: Proportions are with respect to all trials within a given condition. In each panel the upper row of plots shows data from the ASD group, the lower row from the Control group. Each column represents a different search condition (see plot titles). Within each plot, the dots are individual proportions; the bar to the right of each plot shows the mean over subjects and standard error. Horizontal lines indicate min and max values within condition and group. Numbers on the x-axis are individual subject IDs. A: Proportion of trials with RT_{report} < 100 ms. B: Proportion of non-arrival trials. C: Proportion of bad trials. D: Proportion of incorrect trials. E: Proportion of valid trials.

Comparing proportions of valid trials among all trials in a condition

The mean proportion of valid trials among all trials within a condition were around 90±4% in both groups for all conditions (N-search: *M*(ASD) = 91%, *M*(Control) = 94%; RN-search: *M*(ASD) = 91%, *M*(Control) = 93%; X-search: *M*(ASD) = 87%, *M*(Control) = 89%; SX-search: *M*(ASD) = 90%, *M*(Control) = 93%). The groups did not differ with respect to this proportion in any condition (N-search: *t*(34) = -1.53, *p* = .14; RN-search: *t*(34) = -0.59, *p* = .56; X-search: *t*(34) = -0.67, *p* = .50; SX-search: *t*(34) = -1.43, *p* = .16).

Initial trials—proportions of valid trials and of incorrect trials—individual and means

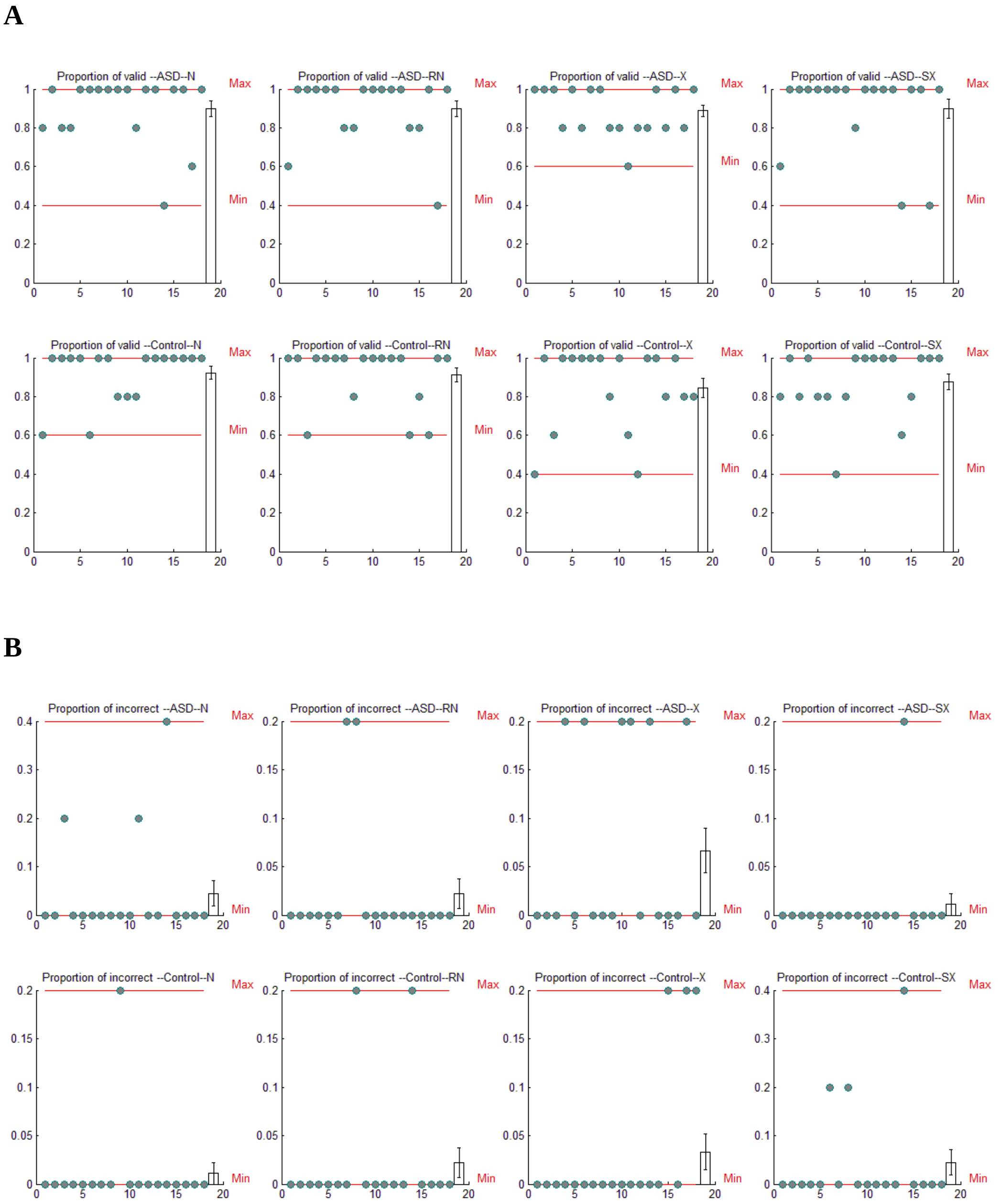


Figure C2. Proportions of valid and of incorrect trials within the first five trials of a given condition. Upper row of plots: ASD group, lower row: Control group. Each column of plots represents a different search condition (see plot titles). Within each plot, the dots are individual proportions; the bar to the right of each plot shows the mean over subjects and standard error. Horizontal lines indicate min and max values within condition and group. Numbers on the x-axis are individual subject IDs. A: Proportion of valid trials. B: Proportion of incorrect trials.

Appendix D: Varying gaze criteria

Table D1 *Explanation for choices of gaze arrival criteria.*

Gaze criteria (pixels)	Explanation
17	0.25 * default criterion
33	Value used by Zhaoping and Guyader (2007)
46	Default criterion; 0.7 * distance between items on regular grid (hereafter: grid size)
60	Value used by Zhaoping and Frith (Exp. 1, 2011) for 12x16 items
66	1 * grid size
79	1.2 * grid size; same fraction as used by Zhaoping and Frith (Exp. 1, 2011)
99	1.5 * grid size
120	Value used by Zhaoping and Frith (Exp. 2, 2011)

Appendix E: Comparing well-tracked and insufficiently tracked participants

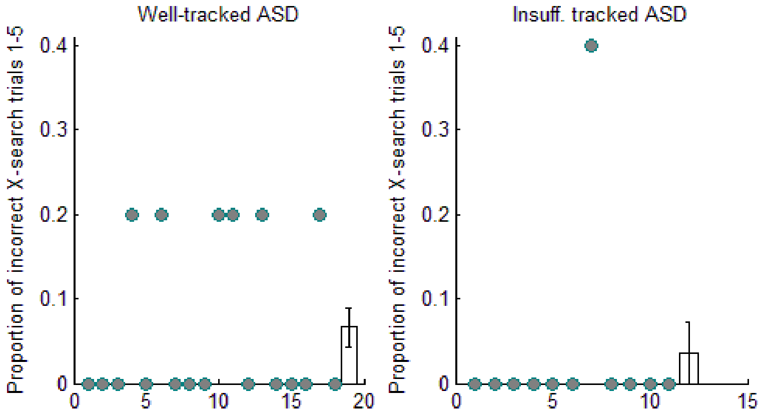


Figure E1. Proportion of incorrect initial X-search trials compared between well-tracked and insufficiently tracked ASD participants. Left panel: Well-tracked ASD subgroup, right panel: insufficiently tracked ASD subgroup. Bars to the right show mean over subjects with standard error. Dots are individual error rates (proportion of incorrect trials out of first 5 trials). Numbers on the x-axis stand for individual subjects of each subgroup.

Appendix F: Figure of all RTs over several trial intervals of well-tracked participants